DYNAMICS READING COURSE SUMMER 2020

SUXUAN CHEN CSUXUAN@GMAIL.COM
JAMES MARSHALL REBER JAMESMARSHALLREBER@GMAIL.COM
HAO-TONG YAN HAOTONGYAN23@GMAIL.COM

Contents

Most of the problems are derived from problems Dr. Gogolev sent us. Each problem should be signed with who did it and with references (if needed). Any references to a proposition/theorem/etc that are not linked and not explicitly stated are implied to be from [1].

As far as reading goes, we followed [1]. The focus was on part 1. Sections covered include 0.3, 0.4, all of chapter 1, 2.1, 2.2, 2.4, 2.5, 3.1 (just the part on topological entropy), 3.2, 3.3, 4.1. Typos which are not in the proofs themselves are most likely by James.

1. Preliminaries

Problems in this section are from the first two handwritten homeworks.

Problem 1 (James)**.** Let $R_\alpha : S^1 \to S^1$ denote the circle rotation $R_\alpha(x) := x + \alpha \pmod 1$. Prove that the special flow over R_α is smoothly conjugate to a linear flow on $\mathbb{T}^2 = S^1 \times S^1$.

Proof. Here, I'm interpreting it as a suspension flow (since we don't have a roof function). We examine the suspension manifold

$$M_{R_\alpha} = [0,1] \times [0,1]/\sim,$$

where $(0,x) \sim (1,x)$ for all x and $(x,1) \sim (x+\alpha \pmod 1, 0)$. The special flow (or here the suspension flow) is then the flow $\sigma^t_{R_\alpha}$ on the M_{R_α} determined by the vertical vector field $\frac{\partial}{\partial t}$. Note that R_α is a diffeomorphism, so we have that M_{R_α} is diffeomorphic to the torus via $h : M_{R_\alpha} \to \mathbb{T}^2$ given by affinely translating $(x,1)$ to $(x+\alpha \pmod 1, 1)$ and linearly interpolating the values (x,y) between $(x,0)$ and $(x+\alpha \pmod 1, 1)$ for $0 < y < 1$. In essence, we are just twisting the torus. Doing so, we see that the vertical flow is now shifted to a linear flow on the torus, and since affine translations are diffeomorphisms, we see that R_α is smoothly conjugate to a linear flow. □

Consider a measure space $(X, \mathcal{M}, \mu)$. We say that a map $T : X \to X$ is *measure preserving* if $\mu(T^{-1}(A)) = \mu(A)$. We say that T is *ergodic* with respect to the measure μ if the following *equivalent* conditions hold:

(1) For every $E \in \mathcal{M}$ with $T^{-1}(E) = E$, either $\mu(E) = 0$ or $\mu(E) = 1$.

(2) For every $E \in \mathcal{M}$ with $\mu(T^{-1}(E)\Delta E) = 0$, we have either $\mu(E) = 0$ or $\mu(E) = 1$.
(3) For every $E \in \mathcal{M}$ with $\mu(E) > 0$, we have
$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(E)\right) = 1.$$
(4) For every two sets E and H of positive measure, there exists an $n > 0$ such that $\mu((T^{-n}(E))\cap H) > 0$.
(5) For every measurable function $f : X \to \mathbb{R}$ with $f \circ T = f$ almost everywhere, we have that f is constant almost everywhere.
(6) (When X compact?) For every $f \in L^2(X, \mu)$ with $f \circ T = f$ almost everywhere, we have f is constant almost everywhere.

For the following exercise, the measure is Lebesgue.

Problem 2 (James)**.** Prove that R_α is ergodic using Fourier Analysis.

Proof. We prove ergodicity using property (6), since $\mathbb{R}/\mathbb{Z}$ compact. Let $f \in L^2(\mathbb{R}/\mathbb{Z}, \lambda)$, where λ is Lebesgue measure. Assume that $f \circ R_\alpha = f$. Since f is in L^2, we have that it is almost everywhere equal to it's Fourier series,
$$f(x) = \sum_{-\infty}^{\infty} a_n e^{2\pi i n x},$$
so
$$f \circ R_\alpha(x) = \sum_{-\infty}^{\infty} a_n e^{2\pi i n(x+\alpha)} = \sum_{-\infty}^{\infty} a_n e^{2\pi i n\alpha} e^{2\pi i n x}.$$
We have that these Fourier series are equal almost everywhere, so we must have the coefficients are equal almost everywhere. Hence,
$$a_n = a_n e^{2\pi i n\alpha} \leftrightarrow a_n(1 - e^{2\pi i n\alpha}) = 0.$$
Notice that, since α is irrational, we have that $1 - e^{2\pi i n\alpha} \neq 0$ for $n \neq 0$, hence $a_n = 0$ for $n \neq 0$. But this then forces f to be constant. □

Problem 3 (James)**.** If $\alpha \notin \mathbb{Q}$, prove that $T(x, y) = (x + \alpha, x + y)$ on $\mathbb{T}^2$ is ergodic.

Proof. We first need to check that T preserves measure with respect to Lebesgue measure. We can equivalently show that, for all $f : \mathbb{T}^2 \to \mathbb{R}$ which are integrable, we have
$$\int f(x, y)d(x \times y) = \int f \circ T d(x \times y).$$
Notice
$$\int f \circ T(x, y)d(x \times y) = \int f(x + \alpha, x + y)d(x \times y).$$
Fubini/Tonelli applies to give
$$\iint_{\mathbb{T}^2} f(x + \alpha, x + y)dydx.$$
Let $z = x + y$, $dz = dy$, we see that
$$\iint_{\mathbb{T}^2} f(x + \alpha, x + y)dydx = \iint_{\mathbb{T}^2} f(x + \alpha, z)dzdx,$$
and doing a change of variables $u = x + \alpha$, we have
$$\iint_{\mathbb{T}^2} f(u, z)dzdu = \int_{\mathbb{T}^2} f(x, y)d(x \times y).$$
Thus, it is measure preserving.

Next, we need to establish ergodicity. Let $f \in L^2(\mathbb{T}^2, \lambda)$ be such that $f \circ T = f$ almost everywhere. The goal is to establish that f is constant almost everywhere. We again use Fourier series. On the torus, we note that

$$f(x,y) = \sum_{n \in \mathbb{Z}^2} a_n e^{2\pi i n \cdot (x,y)} = \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i n_1 x} e^{2\pi i n_2 y}$$

almost everywhere. Note that

$$\begin{aligned} f \circ T(x,y) = f(x+\alpha, x+y) &= \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i n_1 (x+\alpha)} e^{2\pi i n_2 (x+y)} \\ &= \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i (n_1+n_2) x} e^{2\pi i n_1 \alpha} e^{2\pi i n_2 y}. \end{aligned}$$

Doing a shift in n_1 (since it ranges over all integers anyways), we rewrite the series to get

$$\sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1-n_2,n_2)} e^{2\pi i n_1 x} e^{2\pi i (n_1-n_2)\alpha} e^{2\pi i n_2 y}.$$

The same trick as before applies. The coefficients of these series must be equal, so we have

$$a_{(n_1,n_2)} = a_{(n_1-n_2,n_2)} e^{2\pi i (n_1-n_2)\alpha}.$$

Notice that we have $|a_{(n_1,n_2)}| = |a_{(n_1-n_2,n_2)}| = |a_{(n_1-2n_2,n_2)}| = \cdots = |a_{(n_1-kn_2,n_2)}| = \cdots$, where k is an integer. Riemann-Lebesgue forces $a_{(n_1,n_2)} = 0$ if $n_2 \neq 0$. Thus, it suffices to examine the case $n_2 = 0$. Here, we have

$$a_{(n_1,0)} = a_{(n_1,0)} e^{2\pi i n_1 \alpha}.$$

Notice that this is the same as

$$a_{(n_1,0)} \left[1 - e^{2\pi i n_1 \alpha}\right] = 0,$$

so either $a_{(n_1,0)} = 0$ or $e^{2\pi i n_1 \alpha} = 1$. Like before, α is irrational, so this implies that for no non-zero integer n_1 we have $e^{2\pi i n_1 \alpha} = 1$, so $a_{(n_1,0)} = 0$. Hence, the function is constant almost everywhere. □

Remark. The above can be refined to if and only if.

Let $\pi \in \mathrm{Sym}(\{0, \ldots, n\})$ (i.e. a permutation of $\{1, \ldots, n\}$), $v = (v_1, \ldots, v_n)$ a vector in the interior of the unit simplex such that $v_i > 0$ and $\sum v_i = 1$, and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$ a vector with $\epsilon_i = \pm 1$. Let $u_0 = 0$, $u_i = v_1 + \cdots + v_i$ for $i = 1, \ldots, n$, and $\Delta_i = (u_{i-1}, u_i)$ for $i = 1, \ldots, n$. The interval exchange information (denoted IET) is a map $I_{v,\pi,\epsilon} : [0,1] \to [0,1]$ such that it is continuous and Lebesgue measure preserving on every interval Δ_i. The way it works is that it rearranges the intervals Δ_i based on π and either reverses or preserves orientation depending on the sign of ϵ_i.

Problem 4 (Katok 4.1.4, James)**.** Let (X, μ) be a Lebesgue space, $A \subset X$ a measurable set with $\mu(A) > 0$. Let $T : X \to X$ be a measure-preserving transformation, and μ_A the conditional measure defined by

$$\mu_A(B) = \frac{\mu(B \cap A)}{\mu(A)}.$$

For $x \in A$, let $n(x) := \min\{n \in \mathbb{Z}_{\geq 1} : T^n(x) \in A\}$. Prove that the formula $T_A(x) := T^{n(x)}(x)$ defines a transformation of A which preserves the measure μ_A. The map T_A is called the first return map induced by T on the set A.

Proof. Recall the Poincare Recurrence Theorem.

Theorem. Let T be a measure preserving transformation of a Lebesgue space (X, μ) and let $A \subset X$ be a measurable set. Then for any $N \in \mathbb{N}$,

$$\mu(\{x \in A : \{T^n(x)\}_{n \geq N} \subset X \setminus A\}) = 0.$$

We need to show that T_A defines a transformation of A, and this map preserves the measure μ_A. Since T is a measure preserving transformation of a Lebesgue space, Poincare Recurrence says that, for almost every $x \in A$, there exists an n so that $T^n(x) \in A$. So we can relabel the set A to exclude a set of measure zero so that T_A maps A to A (I believe this is fine, but I haven't read the entirety of the chapter to know for sure).

Now, we need to establish that T_A is measure preserving. Recall (**Folland Exercise 1.24**) that we can define a σ-algebra for A by restricting; i.e., define

$$\Sigma_A = \{B \cap A : B \in \Sigma\}.$$

Then the conditional measure μ_A is indeed a measure for this space. Let $B \subset A$ be a measurable set. The goal is to show that

$$\mu_A(T_A^{-1}(B)) = \mu_A(B) = \frac{\mu(B)}{\mu(A)}.$$

The map $n : X \to \mathbb{N}$ defined by $n(x) = \min\{n \in \mathbb{Z}_{\geq 1} : T^n(x) \in A\}$ is measurable. Hence, define

$$A_k = \{x \in X : n(x) = k\}.$$

We have that the A_k partition A, so that

$$A = \bigsqcup_{k \geq 1} A_k.$$

If $B \subset A$ a measurable set, we have

$$B = B \cap A = \bigsqcup_{k \geq 1} (B \cap A_k).$$

Hence,

$$T_A^{-1}(B) = \bigsqcup_{k \geq 1} A_k \cap T^{-k}(B).$$

Notice that

$$\mu(T_A^{-1}(B)) = \mu\left(\bigsqcup_{k \geq 1} A_k \cap T^{-k}(B)\right) = \sum_{k \geq 1} \mu(A_k \cap T^{-k}(B)).$$

It now suffices to show that $\mu(B)$ is this sum. Let $F_0 = A$, $F_1 = T^{-1}(A) \setminus A$, and recursively $F_j = T^{-j+1}(A_{j-1}) \setminus A$. We then note that

$$F_j = \{x \in X : T^j(x) \in A, T^k(x) \notin A \text{ for } 0 \leq k < j\}.$$

We have the identity

$$\mu(F_j) = \mu(F_{j+1}) + \mu(A_{j+1}),$$

since

$$\begin{aligned} T^{-1}(F_j) = \{x \in X : T(x) \in F_j\} = \{x \in X : T^j(T(x)) \in A, T^k(T(x)) \notin A \text{ for } 0 \leq k < j\} \\ = \{x \in A : T^{j+1}(x) \in A, T^k(x) \notin A \text{ for } 1 \leq k < j+1\} \\ \sqcup \{x \in A^c : T^{j+1}(x) \in A, T^k(x) \notin A \text{ for } 1 \leq k < j+1\} \\ = A_{j+1} \sqcup F_{j+1}, \end{aligned}$$

so

$$\mu(T^{-1}(F_j)) = \mu(F_j) = \mu(A_{j+1}) + \mu(F_{j+1}).$$

Hence,

$$\begin{aligned}\mu(B) = \mu(F_0 \cap B) &= \mu(F_1 \cap T^{-1}(B)) + \mu(A_1 \cap T^{-1}(B)) \\ &= \mu(F_2 \cap T^{-2}(B)) + \mu(A_2 \cap T^{-2}(B)) + \mu(A_1 \cap T^{-1}(B)) \\ &= \cdots = \sum_{k\geq 1} \mu(A_k \cap T^{-k}(B)),\end{aligned}$$

since

$$\begin{aligned}F_k \cap T^{-k}(B) &= \{x \in X : T^k(x) \in B, T^j(x) \notin A \text{ for } 0 \leq j < k\} \\ &= (F_{k+1} \cap T^{-k-1}(B)) \sqcup (A_{k+1} \cap T^{-k-1}(B)).\end{aligned}$$

So

$$\mu_A(T_A^{-1}(B)) = \frac{\mu(T_A^{-1}(B))}{\mu(A)} = \frac{\mu(B)}{\mu(A)} = \mu_A(B).$$

This applies for all $B \subset A$ measurable, so T_A is measure preserving. □

Problem 5 (Katok 4.2.10, James)**.** Let $0 < \alpha < \beta < 1$ and consider the following piecewise-continuous transformation $I_{\alpha,\beta}$ of the interval $[0, 1)$ to itself:

$$I_{\alpha,\beta}(x) := \begin{cases} x + 1 - \alpha, \text{ if } 0 \leq x < \alpha, \\ x - \alpha + 1 - \beta, \text{ if } \alpha \leq x < \beta \\ x - \beta, \text{ if } \beta \leq x < 1. \end{cases}$$

Prove that $I_{\alpha,\beta}$ is an injective map which preserves Lebesgue measure. Prove that $I_{\alpha,\beta}$ is ergodic with respect to Lebesgue measure if and only if $\beta/(1 + \beta - \alpha)$ is irrational.

Proof. The map is injective, since it's just a piecewise linear map switching the intervals $[\beta, 1)$ and $[0, \alpha)$. The open intervals generate the σ-algebra, so checking on those is sufficient. It's clear that it preserves the length of this interval, since it just rearranges parts of it.

To prove the latter part, we need to find an injective Lebesgue measure preserving correspondence between $I_{\alpha,\beta}$ and the first return map on a certain rotation on a certain interval $I \subset S^1 = [0, 1]/\sim$. Consider the interval

$$\left[0, \frac{1}{1 + \beta - \alpha}\right)$$

and the associated rotation

$$R_{\frac{1-\alpha}{1+\beta-\alpha}} : S^1 \to S^1.$$

Then if we scale our points α and β, we have

$$\alpha' = \frac{\alpha}{1 + \beta - \alpha},$$

$$\beta' = \frac{\beta}{1 + \beta - \alpha}.$$

We then claim that this switches the intervals $[0, \alpha')$ and $[\beta', \frac{\beta-\alpha}{1+\beta-\alpha})$ and leaves the inner interval in the center. Note that $[0, \alpha') \mapsto \left[\frac{1-\alpha}{1+\beta-\alpha}, \frac{1}{1+\beta-\alpha}\right)$, so the first interval indeed maps to the end. For the inner interval, we see that

$$\alpha' + \frac{1 - \alpha}{1 + \beta - \alpha} = \frac{2 - \alpha}{1 + \beta - \alpha}.$$

Notice $0 < \alpha < \beta < 1$, so

$$1 < 1 + \alpha < 1 + \beta < 2, 1 < 1 + \beta - \alpha < 2 - \alpha,$$

so

$$\frac{1}{2-\alpha} < \frac{1}{1+\beta-\alpha} < 1,$$

and

$$0 < \alpha < 1, -1 < -\alpha < 0, 1 < 2-\alpha < 2,$$

so

$$1 < \frac{2-\alpha}{1+\beta-\alpha} < 2,$$

and hence subtracting 1 gives us $\frac{1-\beta}{1+\beta-\alpha}$, which is in the desired interval. We see then that going up to β', we have that the second interval maps to $\left[\frac{1-\beta}{1+\beta-\alpha}, \frac{1-\alpha}{1+\beta-\alpha}\right)$. The same calculation shows that $\beta' \mapsto 0$ and $\frac{1}{1+\beta-\alpha} \mapsto \frac{\beta-1}{1+\beta-\alpha}$. So scaling this up by $(1+\beta-\alpha)$, we get the desired IET. Hence, the two dynamical systems are related by the scaling map, which we see is a measure invariant mapping which is injective, so we get that $I_{\alpha,\beta}$ is induced by a circle rotation.

Next, we claim the induced rotation is ergodic iff $\beta/(1+\beta-\alpha)$ is irrational. □

Problem 6 (James)**.** Consider the interval exchange information $F : [0,1) \to [0,1)$ of 3 intervals: for fixed $a, b \in [0,1)$, we have $\Delta_1 = [0,a)$, $\Delta_2 = [a,b)$, and $\Delta_3 = [b,1)$. F is then the IET determined by $\pi = (13)$. Show that F is induced by a circle rotation. Under which conditions on the lengths of Δ_i do we have that F is minimal?

Proof. I paint most of the picture in a blog post on my website. See here. □

Problem 7 (James)**.** Define

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

by

$$T(x,y) = (2x+y, x+y) \pmod 1.$$

(1) Prove that T is ergodic with respect to Lebesgue measure.
(2) Prove that T is topologically mixing.
(3) Prove that the periodic points of T are dense in $\mathbb{T}^2$.
(4) Let $T \in \mathrm{SL}(d,\mathbb{Z})$. Then T induces $T : \mathbb{T}^d \to \mathbb{T}^d$. Assume λ is an eigenvalue of T and $\lambda^m = 1$ for some m. Prove that T is not ergodic with respect to Lebesgue measure.

Remark. This is similar to future problems, but I included it as a reference since the approach here may be different.

Before proceeding, we give a useful lemma on measure preserving transformations.

Lemma. Assume (X, Σ, μ) is a σ-finite measure space. A transformation $T : (X,\Sigma,\mu) \to (X,\Sigma,\mu)$ is μ measure preserving iff for all $f : X \to \mathbb{R}$ with f integrable, we have

$$\int f d\mu = \int f \circ T d\mu.$$

Proof. $(\implies)$: Assume that T is measure preserving. Then for $A \in \Sigma$ measurable, we have that $\mu(T^{-1}(A)) = \mu(A)$. In other words,

$$\mu(A) = \int \chi_A d\mu = \mu(T^{-1}(A)) = \int \chi_{T^{-1}(A)} d\mu = \int \chi_A \circ T d\mu.$$

So for all characteristic functions, we have that it holds. Let φ be a simple function, i.e.

$$\varphi = \sum a_n \chi_{A_n}.$$

Then

$$\begin{aligned}\int \varphi d\mu = \int \left(\sum a_n \chi_{A_n}\right) d\mu = \sum a_n \int \chi_{A_n} d\mu = \sum a_n \mu(A_n) \\ = \sum a_n \mu(T^{-1}(A_n)) = \sum a_n \int \chi_{A_n} \circ T d\mu \\ = \int \sum a_n \chi_{A_n} \circ T d\mu = \int \varphi \circ T d\mu.\end{aligned}$$

It holds for simple functions then. Let $f : X \to \mathbb{R}_{\geq 0}$ be such that $f \in L^+(X, \Sigma, \mu)$. Then we have that we can construct a monotone sequence of simple functions φ_n so that $\varphi_n \nearrow f$. The monotone convergence theorem implies that

$$\int f d\mu = \int \lim_{n\to\infty} \varphi_n d\mu = \lim_{n\to\infty} \int \varphi_n d\mu = \lim_{n\to\infty} \int \varphi_n \circ T d\mu = \int \lim_{n\to\infty} \varphi_n \circ T d\mu = \int f \circ T d\mu.$$

For $f \in L^1(X, \Sigma, \mu)$, write $f = f_+ - f_-$, so it is a linear combination of its positive and negative parts. Then

$$\int f d\mu = \int f_+ d\mu - \int f_- d\mu.$$

Since $f_+, f_- \in L^+(X, \Sigma, \mu)$, we have that

$$\int f_+ d\mu = \int f_+ \circ T d\mu,$$

$$\int f_- d\mu = \int f_- \circ T d\mu,$$

so

$$\int f d\mu = \int f_+ \circ T d\mu - \int f_- \circ T d\mu = \int f \circ T d\mu.$$

Hence, it holds for all integrable functions.
($\Longleftarrow$): Assume the property holds for all integrable functions. Let $A \in \Sigma$ be a measurable set with finite measure. Then χ_A is integrable, so

$$\mu(A) = \int \chi_A d\mu = \int \chi_A \circ T d\mu = \int \chi_{T^{-1}(A)} d\mu = \mu(T^{-1}(A)).$$

So it holds for all sets with finite measure. Assume A has infinite measure. Then we can write $A = \bigcup A_j$, where $\mu(A_j) < \infty$. In particular, we can make this sequence increasing, so that

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Note that the monotone convergence theorem says

$$\mu(A) = \int \chi_A d\mu = \lim_{n\to\infty} \int \chi_{A_n} d\mu = \lim_{n\to\infty} \int \chi_{A_n} \circ T d\mu = \int \chi_A \circ T d\mu = \mu(T^{-1}(A)).$$

□

Remark. One question is whether the σ-finite property is necessary. I couldn't find a good reference for this lemma (even though it's used in different places) so I don't know whether it was. My guess is it is not necessary.

Proof.

(1) By the lemma above, it suffices to check that for $f:\mathbb{T}^2\to\mathbb{R}$ integrable, we have

$$\int f(x,y)d(x\times y)=\int f\circ T(x,y)d(x\times y).$$

Lebesgue measure is σ-finite, so it is fine to use the lemma. Note that by definition of T, we have

$$\int_{\mathbb{T}^2} f\circ T(x,y)d(x\times y)=\int_{\mathbb{T}^2} f(2x+y,x+y)d(x\times y).$$

Note that Fubini-Tonelli works here, so it is fine to iterate the integral to get

$$\iint_{\mathbb{T}^2} f(2x+y,x+y)dxdy.$$

A change of variables $u=x+y$, $z=x+u$ gives $du=dy$, $dz=dx$, and so we get the integral is equal to

$$\iint_{\mathbb{T}^2} f(z,u)dzdu=\int_{\mathbb{T}^2} f(x,y)d(x\times y).$$

Hence, T is measure preserving.

We now need to show that T is ergodic with respect to Lebesgue measure. We invoke the L^2 definition of ergodic, so T is ergodic iff for all $f\in L^2(\mathbb{T}^2,\lambda)$, we have that $f\circ T=f$ implies f is constant almost everywhere. We have that f in $L^2(\mathbb{T}^2,\lambda)$ implies that we can express it almost everywhere in terms of its Fourier series,

$$f(x,y)=\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)}e^{2\pi i(n_1,n_2)\cdot(x,y)}$$

so

$$\begin{aligned}f\circ T(x,y)=\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)}e^{2\pi i(n_1,n_2)\cdot(2x+y,x+y)}&=\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)}e^{2\pi i n_1(2x+y)}e^{2\pi i n_2(x+y)}\\&=\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)}e^{2\pi i x(2n_1+n_2)}e^{2\pi i y(n_1+n_2)}.\end{aligned}$$

Let $k=n_1+n_2$, then $n_2=k-n_1$, so the series can be expressed as

$$\sum_{(n_1,k)\in\mathbb{Z}^2} a_{(n_1,k-n_1)}e^{2\pi i x(n_1+k)}e^{2\pi i yk}.$$

Letting $u=n_1+k$, we have $n_1=u-k$, so

$$\sum_{(u,k)\in\mathbb{Z}^2} a_{(u-k,2k-u)}e^{2\pi i xu}e^{2\pi i yk}.$$

Relabeling u and k, we have that the series is equivalent to

$$\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1-n_2,2n_2-n_1)}e^{2\pi i xn_1}e^{2\pi i yn_2},$$

so since these are equal (almost everywhere), we have that the Fourier series agree, so

$$a_{(n_1,n_2)}=a_{(n_1-n_2,2n_2-n_1)}.$$

Fixing $(n_1,n_2)\in\mathbb{Z}^2$ non-trivial, we iterate to see that

$$a_{(n_1,n_2)}=a_{(n_1-n_2,2n_2-n_1)}=a_{(2n_1-3n_2,5n_2-3n_1)}=\cdots.$$

We see its equal to infinitely many distinct coefficients, so Riemann-Lebesgue says that it must be 0. Hence, it's constant almost everywhere, so T is ergodic.

(2) A dynamical system is said to be *topologically mixing* if, for every two non-empty open sets U and V in X, there exists an N sufficiently large so that $f^n(U) \cap V \neq \varnothing$ for all $n \geq N$.

A measure preserving transformation $T : (X, \Sigma, \mu) \to (X, \Sigma, \mu)$ is *mixing* (with respect to the measure μ) if for any two measurable $A, B \subset X$, we have

$$\mu(T^{-n}(A) \cap B) \to \mu(A) \cdot \mu(B) \text{ as } n \to \infty.$$

It follows that if a measure preserving transformation is mixing, then it is topologically mixing (so long as the open subsets have positive measure). In this case, Lebesgue measure is such that non-empty open sets have positive measure (take a point in the set, then you can find an open interval contained in the open set, and the open interval has positive measure). Hence, for $U, V \subset \mathbb{T}^2$ open and non-empty, if T is mixing with respect to the Lebesgue measure, then $\mu(T^{-n}(A) \cap B) > 0$ for n large enough, implying that $T^{-n}(A) \cap B \neq \varnothing$ for n large enough.

Let H be a Hilbert space. A family of functions $\mathcal{F} \subset H$ is said to be *complete* if there is no function $f \in H$ so that $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{F}$. In other words, $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{F}$ implies $f = 0$.

We have the following proposition from Katok.

Proposition 1. Consider a measure space (X, Σ, μ). T is mixing (with respect to the measure) if and only if for any complete system $\mathcal{F} \subset L^2(X, \mu)$ of functions and any $f, g \in \mathcal{F}$, we have

$$\int f(T^n(x))\overline{g}(x) d\mu \to \left(\int f d\mu \right) \cdot \left(\int \overline{g} d\mu \right) \text{ as } n \to \infty. \tag{1}$$

Proof. **Step 1:** We show that if condition (1) holds for a complete system $\mathcal{F}$, then it holds for linear combinations of elements in $\mathcal{F}$.

To see this, note that the inner product in L^2 is given by

$$\langle f, g \rangle = \int f \overline{g} d\mu.$$

So we rewrite condition (1) as

$$\langle f \circ T^n, g \rangle \to \left(\int f d\mu \right) (\overline{g} d\mu). \tag{2}$$

Now, if condition (2) holds for all of the complete system, then it holds for linear combinations, since

$$\langle (\alpha_1 f_1 + \alpha_2 f_2) \circ T^n, (\beta_1 g_1 + \beta_2 g_2) \rangle = \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_i \overline{\beta_j} \langle f_i \circ T^n, g_j \rangle,$$

and so

$$\begin{aligned}
\lim_{n\to\infty}\langle(\alpha_1 f_1+\alpha_2 f_2)\circ T^n,(\beta_1 g_1+\beta_2 g_2)\rangle &= \lim_{n\to\infty}\sum_{i=1}^{2}\sum_{j=1}^{2}\alpha_i\overline{\beta_j}\langle f_i\circ T^n, g_j\rangle\\
&= \sum_{i=1}^{2}\sum_{j=1}^{2}\alpha_i\overline{\beta_j}\lim_{n\to\infty}\langle f_i\circ T^n, g_j\rangle\\
&= \sum_{i,j=1}^{2}\alpha_i\overline{\beta_j}\left(\int f_i d\mu\right)\cdot\left(\int \overline{g_j}d\mu\right)\\
&= \left(\int \alpha_1 f_1+\alpha f_2 d\mu\right)\cdot\left(\int \overline{\beta_1 g_1+\beta_2 g_2}d\mu\right)
\end{aligned}$$

using the linearity of integrals.

Step 2: It holds for linear combinations of elements in $\mathcal{F}$, so in particular it holds for a dense subset $L(\mathcal{F})\subset L^2(X,\mu)$. This follows, since the span of $\mathcal{F}$ will contain an orthonormal basis, so will be dense (with respect to the topology generated by the norm $\|f\|=\sqrt{\langle f,f\rangle}$).

Step 3: We now wish to show condition (1) or (2) holds for all $f,g\in L^2(X,\mu)$. Since $L(\mathcal{F})$ is dense, for $\epsilon>0$ and $f,g\in L^2(x,\mu)$ we can find $f',g'\in L(\mathcal{F})$ so that $\|f-f'\|<\epsilon$, $\|g-g'\|<\epsilon$. We see that (using the Schwarz inequality)

$$\begin{aligned}
&|\langle f\circ T^n,g\rangle-\langle f,g\rangle|\\
&=\Bigg|\langle f\circ T^n,g-g'\rangle+\langle(f-f')\circ T^n,g\rangle+\langle f'\circ T^n,g'\rangle\\
&-\left(\int f'd\mu\right)\left(\int\overline{g'}d\mu\right)+\left(\int f'd\mu\right)\left(\int\overline{g'-g}d\mu\right)+\left(\int(f'-f)d\mu\right)\left(\int\overline{g}d\mu\right)\Bigg|\\
&\le\|f\circ T^n\|\|g-g'\|+\|(f-f')\circ T^n\|\|g'\|\\
&+\left|\int f'(T^n(x))\overline{g'}(x)d\mu-\left(\int f'd\mu\right)\left(\int\overline{g'}d\mu\right)\right|+\left|\int f'd\mu\right|\cdot\|g-g'\|+\|f-f'\|\cdot\left|\int\overline{g}d\mu\right|\\
&\le\left|\int f'(T^n(x))\overline{g'}(x)d\mu-\left(\int f'd\mu\right)\left(\int\overline{g'}d\mu\right)\right|+\epsilon\left(\|f\|+1+\left|\int f'd\mu\right|+\left|\int\overline{g}d\mu\right|\right).
\end{aligned}$$

If we take $\epsilon\to0$, we get the desired result, which is that condition (2) holds.

Step 4: We finish the proof. If T is mixing, then we note that the set $\mathcal{F}=\{\chi_A : A \text{ is measurable}\}$ forms a complete system, so by **Step 3** we have that condition (1) holds for all $L^2(X,\mu)$. If condition (1) holds on a complete system, then by **Step 3** we see that it holds on all of $L^2(X,\mu)$, so in particular it holds on $f=\chi_A$ and $g=\chi_B$ for $A,B\subset X$ measurable, and so T is mixing. □

We now show that T is topologically mixing by showing that T is mixing with respect to Lebesgue measure, which means showing that condition (1) holds on a complete system. The system $\mathcal{F}=\{\exp(2\pi i(m,n)\cdot(x,y)) : (m,n)\in\mathbb{Z}^2\}$ forms a complete system (by Stone-Weierstrass), so we just need to show condition (1) holds on this system. Let

$$\chi_{(m,n)}(x,y)=\exp(2\pi i(m,n)\cdot(x,y)).$$

So we just need to show that

$$\int\chi_{(m,n)}(T^l(x))\overline{\chi_{(j,k)}}d\mu\to\left(\int\chi_{(m,n)}d\mu\right)\cdot\left(\int\chi_{(j,k)}d\mu\right).$$

If $m = n = j = k = 0$, we get that $\chi_{(0,0)} = 1$, so both the left and right sides are 1. Assume $(m, n) \neq (0, 0)$. We see that the right hand side of the equation is 0, so it suffices to show that the left hand side converges to 0. Notice that

$$\chi_{(m,n)}(T(x)) = \exp(2\pi i(m, n) \cdot (2x + y, x + y)).$$

We can correspond to the action of T on $\mathbb{T}^2$ the matrix

$$\widehat{T} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

so that

$$\chi_{(m,n)}(T^N(x, y)) = \exp(2\pi i(m, n) \cdot T^N(x, y)) = \exp(2\pi i \cdot \widehat{T}^N(m, n) \cdot (x, y)) = \chi_{\widehat{T}^N(m,n)}(x, y)$$

We can now rewrite the left hand side as

$$\int \chi_{\widehat{T}^N(m,n)-(j,k)} d\mu.$$

Now, this is zero so long as $\widehat{T}^N(m, n)-(j, k) \neq 0$, or $\widehat{T}^N(m, n) \neq (j, k)$. Since $\|\widehat{T}^N(m, n)\| \to \infty$, there exists sufficiently large N so that $\widehat{T}^N(m, n) \neq (j, k)$. So we have that the integral converges to 0.

(3) We now want to show that the space of periodic points are dense in $\mathbb{T}^2$. This follows by showing that coordinates with rational components are periodic. Let $x = (p/q, s/q) \in \mathbb{T}^2$, p, q, s are integers. The goal is to show there exists an N so that $T^N(x) = x$. Note that

$$T(x) = \left(\frac{2p + s}{q}, \frac{p + s}{q}\right).$$

So this maps to a rational coordinate with denominator q. There are q ways of creating a coordinate with denominator q in the first component (since $0 \leq p/q < 1$), and q ways of doing it for the second component, so in total there are q^2 ways of doing it. So $\mathcal{O}_T(x)$ has finite order, hence must repeat. So there is a m, n so that $T^n(x) = T^m(x)$. The map T is invertible, since it corresponds to a matrix with determinant non-zero, so if $n \geq m$ we have $T^{n-m}(x) = x$. Hence, it is periodic. Thus, periodic points are dense.

(4) We have that T is a matrix

$$T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix}$$

such that $\det(T) = \pm 1$, $a_{ij} \in \mathbb{Z}$. Notice we have that the transformation corresponds to

$$T(x_1, \ldots, x_d) = \left(\sum a_{1i}x_i, \sum a_{2i}x_i, \ldots, \sum a_{di}x_i\right) \pmod 1.$$

Consider the spectrum $\operatorname{spec}(T) = (\lambda_1, \ldots, \lambda_d)$ (where there may be repeats). Assume that $\lambda_i^m = 1$ for some m. [TODO]

□

2. Topological Dynamics

Problem 8 (Problem 1.1, James, Suxuan)**.** Prove that an irrational circle rotation $R_\alpha : x \mapsto x + \alpha \pmod 1$, $\alpha \notin \mathbb{Q}$, is minimal.

Proof. The dynamical system is minimal iff

$$\bigcap_{x\in X} \overline{\mathcal{O}_{R_\alpha}(x)} = X.$$

Assume for contradiction that it is not minimal. Then there is some $x \in X$ with $A := \overline{\mathcal{O}_{R_\alpha}(x)} \subset X$. Since A is closed, we have that A^c is open, and since the open intervals generate the topology of $\mathbb{R}$ we get that there is an open interval $I \subset A$. We note that for intervals $I = (a, b)$, $R_\alpha(I) = (a+\alpha, b+\alpha)$ (mod 1), so $\lambda(R_\alpha(I)) = \lambda(I)$ – that is, rotations preserve length. If we consider

$$\mathcal{F} = \{I \text{ an interval} : I \subset A\},$$

we get a partial ordering by containment. We note that the union of a chain of intervals is an interval, so Zorn's lemma says there is an interval of maximum length contained in A. Let I be this interval.

First, we note that $R^n_\alpha(I) \subset A^c$. If $R^n_\alpha(I) \subset A$ for some n, then since R_α is invertible we get that $I \subset \overline{\mathcal{O}_{R_\alpha}(x)}$, a contradiction.

Next, since $\{R^n_\alpha(I) : n \in \mathbb{Z}\}$ is a collection of intervals in A^c, we note that they partially cannot overlap. If $R^n_\alpha(I) \cap R^m_\alpha(I) \neq \varnothing$ for some $n \neq m$, then we have that $R^n_\alpha(I) \cup R^m_\alpha(I) \subset A^c$ is an interval with size larger than I, contradicting maximality.

Finally, we claim that $R^n_\alpha(I) \neq R^m_\alpha(I)$ for $n \neq m$. If it did hold, then by the invertibility of R_α we get that $R^k_\alpha(I) = I$ for some integer k, so the endpoint a of I is mapped to a after k iterates, implying that it is a periodic point. Thus, we have $a + k\alpha \equiv a \pmod 1$, $k\alpha$ an integer, and so α is rational, a contradiction.

So $\{R^n_\alpha(I) : n \in \mathbb{Z}\}$ is an infinite collection of disjoint intervals in A^c of equal length. This is a contradiction, since S^1 has finite length. $\square$

Suxuan's Proof. First we show that for given $N \in \mathbb{N}$, there exists $n, m \in \mathbb{N}$ such that $|R^n(0) - R^m(x)| \leq \frac{1}{N}$. Consider the partition $0 < \frac{1}{N} < \frac{2}{N} < \cdots < \frac{N-1}{N} < 1$ of the unit interval $[0, 1]$, then we get N small subintervals. For $n, m \in \mathbb{N}$, $n \neq m$, suppose $R^n(x) = R^m(x)$, then

$$x + n\alpha \equiv x + m\alpha \pmod 1,$$

so there exists $b \in \mathbb{Z}$ such that $n\alpha - m\alpha = b$, and so $\alpha = \frac{b}{n-m}$, which contradicts to the condition $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, hence $R^n(x) \neq R^m(x)$ if $n \neq m$. Then the $N + 1$ points $R(x), R^2(x), \ldots, R^{N+1}(x)$ are $N + 1$ distinct points in the unit interval $[0, 1]$, by pigeonhole principal, there exists $1 \leq n < m \leq N + 1$ such that $R^n(x)$ and $R^m(x)$ are in the same subinterval, so we have

$$|R^n(x) - R^m(x)| \leq \frac{1}{N}.$$

Since $x + n\alpha \equiv R^n(x)$ (mod 1) and $x + m\alpha \equiv R^m(x)$ (mod 1), we then have

$$m\alpha - n\alpha \equiv R^m(x) - R^n(x) \pmod 1.$$

For $k \in \mathbb{N}$, $R^{k(m-n)}(x) - R^{(k-1)(m-n)}(x) \equiv m\alpha - n\alpha$ (mod 1), so

$$R^{k(m-n)}(x) - R^{(k-1)(n-m)}(x) \equiv R^m(x) - R^n(x) \pmod 1.$$

Without loss of generality, we assume $R^m(x) - R^n(x)$ is positive. Then the intervals $[x, R^{(m-n)}(x)]$, $[R^{(m-n)}(x), R^{2(m-n)}(x)]$, ..., form a cover $[0, 1)$, so for given $y \in [0, 1)$, there exists $K \geq 0$ such that $y \in [R^{K(m-n)}(x), R^{(K+1)(m-n)}(x)]$, hence the orbit of x is dense. $\square$

Problem 9 (Problem 1.2, James)**.** Consider a translation $T : \mathbb{T}^2 \to \mathbb{T}^2$ given by $(x, y) \mapsto (x + \alpha, y + \beta)$. Classify all orbits $\overline{\mathcal{O}_T(x, y)}$.

Proof.

Case 1: If α, β are rational numbers, say

$$\alpha = \frac{p}{q}, \qquad \beta = \frac{k}{l}, \qquad p, q, k, l \in \mathbb{Z}, \quad l, q \neq 0,$$

then we see that for all $(x, y) \in \mathbb{T}^2$ we have that $\mathcal{O}_T(x, y)$ is periodic. This follows by noting that if we choose $m = \text{lcm}(q, l)$, then we have

$$T^m(x, y) = (x + m\alpha, y + m\beta) \equiv (x, y) \pmod 1$$

So $\mathcal{O}_T(x, y)$ is a finite collection of points, and hence $\overline{\mathcal{O}_T(x, y)} = \mathcal{O}_T(x, y)$.

Case 2: If α is irrational, β rational, then we can write β in the reduced form

$$\beta = \frac{p}{q}, \qquad p, q \in \mathbb{Z}, \quad q \neq 0.$$

We claim that $\overline{\mathcal{O}_T(x, y)}$ is a union of q disjoint circles. We know that restricting to the second coordinate, we have that the orbit with respect to T will be periodic with period given by q. We see that for $0 \leq j < q$, $j \in \mathbb{Z}$, $n \in \mathbb{Z}$, we have that

$$T^{(j+nq)}(x, y) = (x + (j + nq)\alpha, y + (j + nq)\beta) \equiv (x + (j + nq)\alpha, y + j\beta) \pmod 1.$$

The second coordinate is then fixed for all $n \in \mathbb{Z}$, and we see that in the first coordinate we have a rotation by $(j + nq)\alpha$ which is still irrational. Iterating this and taking its closure, we get that this corresponds to the circle $S^1 \times \{y + j\beta\}$. This is disjoint from all of the other circles as we look over the range of j, so we get q distinct circles. Notice that this partitions all of the orbit, and we use that the closure of the union is the union of the closure to get

$$\overline{\mathcal{O}_T(x, y)} = \bigsqcup_{\substack{0 \leq j < q \\ j \in \mathbb{Z}}} S^1 \times \{y + j\beta\}.$$

Case 3: Now we need to consider the case α, β are irrational but rationally dependent. In other words, there exists $k_1, k_2, k \in \mathbb{Z}$ with

$$k_1\alpha + k_2\beta = k,$$

and at least one k_1, k_2 non-zero. It turns out both must be non-zero, since otherwise this implies that either α or β are rational, which contradicts the original assumption. Choosing $(x, y) \in \mathbb{T}^2$, we see that

$$T(x, y) = (x + \alpha, y + \beta) \pmod 1.$$

We know this may not be topologically transitive by **Katok Proposition 1.3.4** and **Katok Proposition 1.4.1**. As noted in these propositions, it really suffices to look at the orbit of 0. Doing so, we have that

$$(n\alpha, n\beta) = \left(\frac{nk}{k_1} - \frac{k_2}{k_1}n\beta, n\beta\right).$$

Since the y coordinate is such that $y = n\beta$, we see that we have $x = nk/k_1 - (k_2/k_1)y$, or $y = nk/k_2 - (k_1/k_2)x$. So the iterates of 0 lie along these lines. There are finitely many of these lines in $[0, 1]^2$ (varying over different n values modulo 1), and the orbits are dense on these lines by irrationality, so we get a finite union of circles again.

Case 4: Finally we consider the case α, β are irrational and rationally independent. Then as shown in **Katok 1.4.1**, we see that T is minimal, so for all (x, y) we have $\overline{\mathcal{O}_T(x, y)} = \mathbb{T}^2$.

$\square$

Remark. Case 3 was the trickiest, and it relies on **Lemma 5.1.10** in **A First Course in Dynamics** by Hasselblatt and Katok. However, the proof given in the book is incorrect, so I had to mess around with things (see the errata).

Problem 10 (Problem 1.3, James, Hao)**.** Let $T : X \to X$ be a homeomorphism of a compact metric space. Prove that the following are equivalent:

(1) T is minimal.
(2) For all $U \subset X$ open, there exists an n such that

$$\bigcup_{j=-n}^{n} T^j(U) = X.$$

Proof. (1) $\implies$ (2): Let $U \subset X$ be open and let $x \in X$ be arbitrary. Since T is minimal, this implies that for all $x \in X$ we have a $n \in \mathbb{Z}$ with $T^n(x) \in U$. In other words, for all $x \in X$, there exists $n \in \mathbb{Z}$ with $x \in T^n(U)$, so we have that

$$X \subset \bigcup_{j=-\infty}^{\infty} T^j(U).$$

Since f is a homeomorphism, these sets are all open. By compactness there is a finite refinement, so there exists sufficiently large n so that

$$X = \bigcup_{j=-n}^{n} T^j(U).$$

(2) $\implies$ (1): (*Hao's Solution*) First, we show the intersection property (see **Katok Lemma 1.4.2**). Since we have a metric space, there exists a countable base $\{U_n\}_{n=1}^{\infty}$. By property (2), we have that for each U_j there exists an N_j so that

$$X = \bigcup_{n=-N_j}^{N_j} T^n(U_j).$$

Hence for each U_m in the base, we have

$$U_m = \bigcup_{n=-N_j}^{N_j} (T^n(U_j) \cap U_m),$$

so there exists an integer n so that $T^n(U_j) \cap U_m \neq \varnothing$. Now let U and V be nonempty open sets. Since we have a base, we have that there is a U_j, U_m so that $U_j \subset U$, $U_m \subset V$. By what we've just shown, we have $T^n(U) \cap V \neq \varnothing$. The intersection property then holds.

Now fix $x \in X$ and let $U_1^n = B_{1/2^n}(x)$. For each n, we can construct a corresponding countable base so that $\{U_1^n, U_2^n, \ldots\}$ is a countable base. By the interscetion property, we see that there is an integer N_1 so that $f^{N_1}(U_1^0) \cap U_2^0 \neq \varnothing$, so we have a nonempty open set V_1^0 so that $\overline{V_1} \subset U_1^0 \cap f^{-N_1}(U_2^0)$. Similarly, we have that there is an N_2 so that $f^{N_2}(V_1^0) \cap U_3^0 \neq \varnothing$, so we can find V_2^0 open so that $\overline{V_2^0} \subset V_1^0 \cap f^{-N_2}(U_3^0)$. Repeating in this fashion, we have that

$$V^0 = \bigcap \overline{V_n^0}$$

is a closed compact set which is an intersection of nested compact sets, so it is nonempty. Similarly, we construct $\{V_n^m\}$ for each base $\{U_n^m\}$, and let

$$V^m = \bigcap \overline{V_n^m}.$$

Note that the $\{V^m\}$ are closed, compact sets, and since $U_n^0 \subset U_k^0$ for $n > k$, we get that these are nested. Finally, let
$$V = \bigcap V^m.$$
V is an intersection of nested closed compact sets, so is nonempty. We have that $x \in V$ is the only element which can be in here, since the balls are decreasing to x. Hence $x \in V$. Like in **Katok Lemma 1.4.2**, we see that the elements in V^m for each m are such that they have dense orbit, so x must have dense orbit. The choice of x was arbitrary, so the system is minimal.
(2) $\implies$ (1): (*My Solution*) Let $x \in X$ be arbitrary, $U \subset X$ a nonempty open set. By assumption 2, we have
$$X = \bigcup_{n=-N}^{N} T^n(U),$$
so there is some integer m with $|m| \leq N$ so that $x \in T^m(U)$. In other words, we have $T^{-m}(x) \in U$ (since T a homeomorphism). Thus, for each open set U, there exists an integer m so that $T^m(x) \in U$, so the orbit intersects every nonempty open set nontrivially, so the orbit must be dense. □

Problem 11 (Problem 1.6, James)**.** Give an example of a transitive homeomorphism of a compact manifold which is not topologically mixing.

Proof. Consider the rotation map on the circle $f : x \mapsto x + \alpha \pmod 1$. We claim that it is not mixing regardless of what α is. Consider an nonempty proper interval $I \subsetneq S^1$ whose length is sufficiently small, say $m(I) = \epsilon < 1/8$, and set $I = (0, \epsilon)$. Let N be any integer. We wish to show there is an $k > N$ so that $f^k(I) \cap I = \varnothing$. We note that f preserves the length of the interval, so $m(f^k(I)) = \epsilon$. Consider the interval $J = (1/2, \epsilon + 1/2)$. Since $\epsilon < 1/8$, we have that $I \cap J = \varnothing$ and the distance between I and J is $1/2$. Assuming that f is topologically mixing, we can find a sufficiently large integer k so that $f^k(I) \cap I \neq \varnothing$, $f^k(I) \cap J \neq \varnothing$. Since the distance between I and J is $1/2$, the only way this could happen is if $m(f^k(I)) \geq 1/2$, but this contradicts the fact that iterates of f have the same length. Hence, there must be a $k > N$ so that $f^k(I) \cap I = \varnothing$, so that f is not topologically mixing. □

Problem 12 (Problem 1.8, James)**.** Let α be an irrational number and $T : \mathbb{T}^2 \to \mathbb{T}^2$ is given by
$$T(x, y) = (x + \alpha, y + x) \pmod 1.$$
Prove that T is transitive.

Remark. T is, in fact, minimal, but this is harder to show.

Proof. We go by contradiction. We utilize **Katok Corollary 1.4.3**. Assume that T not topologically transitive. There exists two disjoint nonempty open sets U and V which are T-invariant; so $T(U) = U$, $T(V) = V$, $U \cap V = \varnothing$. Let χ_U be a characteristic function. Invariance says
$$\chi_U(T(x, y)) = \chi_{T^{-1}(U)}(x, y) = \chi_U(x, y).$$
We are on a finite measure space, so with respect to Lebesgue measure this is going to be in L^2. It is fine, then, to take Fourier expansions. Doing so, we have
$$\chi_U(x) = \sum_{n \in \mathbb{Z}^2} a_n e^{2\pi i n \cdot x},$$
viewing $x \in \mathbb{R}^2/\mathbb{Z}^2$ now. Since the functions are equal, the Fourier expansions are also equal, so
$$\chi_U(T(x)) = \sum_{n \in \mathbb{Z}^2} a_n e^{2\pi i n \cdot (x+\alpha, x+y)} = \sum_{n \in \mathbb{Z}^2} a_n e^{2\pi i [(n_1 x + n_1 \alpha) + (n_2 x + n_2 y)]}$$
$$= \sum_{n \in \mathbb{Z}^2} a_{(n_1, n_2)} e^{2\pi i (n_1 + n_2) x + n_2 y} e^{2\pi i n_1 \alpha}.$$

Doing a change of variables, we get

$$\chi_U(T(x,y)) = \sum_{n\in\mathbb{Z}^2} a_{(n_1-n_2,n_2)}e^{2\pi i n_1 x}e^{2\pi i n_2 y}e^{2\pi i(n_1-n_2)\alpha}$$
$$= \sum_{n\in\mathbb{Z}^2} a_{(n_1,n_2)}e^{2\pi i n_1 x}e^{2\pi i n_2 y}.$$

Hence, $a_{(n_1-n_2,n_2)}e^{2\pi i(n_1-n_2)\alpha} = a_{(n_1,n_2)}$, so we get infinitely many Fourier coefficients which are equal in magnitude when $n_2 \neq 0$. Riemann-Lebesgue says that these must all be zero. Thus, we need only analyze the case where $n_2 = 0$; here, we have

$$a_{(n_1,0)} = a_{(n_1,0)}e^{2\pi i n_1 \alpha},$$

and the usual argument shows that if α is irrational we must have $n_1 = 0$. This forces χ_U to be a constant function, meaning it must be either zero or one. Since U has positive measure and $V \subset U^c$ has positive measure, we get that this is impossible. So there cannot be two disjoint nonempty open sets U and V which are T-invariant, and this implies that T is topologically transitive. $\square$

Problem 13 (Problem 1.9, James, Suxuan)**.** Show that a factor of a topologically mixing system is also topologically mixing.

Proof. Recall that a system $f : X \to X$ is *topologically mixing* if for all U, V nonempty open subsets of X, there exists an N so that for all $n \geq N$ we have

$$f^n(U) \cap V \neq \varnothing.$$

Recall that a map $g : Y \to Y$ is said to be a *factor* of f if there exists a continuous surjective map $h : X \to Y$ with

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

For fixed $n \geq 1$, we see that

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

so iterating we get

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

The goal is to show that for any oepn $U, V \subset Y$ nonempty, there exists an N so that for all $n \geq N$ with

$$g^n(U) \cap V \neq \varnothing.$$

Since U, V are open, we have that $h^{-1}(U), h^{-1}(V) \subset X$ are open sets. Hence, there exists an N so that for $n \geq N$ we have

$$f^n(h^{-1}(U)) \cap h^{-1}(V) \neq \varnothing.$$

Applying h gives us

$$\varnothing \neq h(f^n(h^{-1}(U)) \cap h^{-1}(V)) \subset h(f^n(h^{-1}(U))) \cap h(h^{-1}(V)) = g^n(U) \cap V.$$

This holds for all $n \geq N$, so we get that g is topologically mixing as well. $\square$

Problem 14 (Problem 1.10, James)**.** Prove that $T : \mathbb{T}^2 \to \mathbb{T}^2$ given by $T(x,y) = (2x+y, x+y)$ is topologically mixing.

Remark. I showed this in a sort of indirect way initially. An easy way to see it is topologically mixing is by noting that, since T is expanding along a (dense) eigenline on the torus, we eventually must have $T^n(U) \cap V \neq \varnothing$ since U is eventually stretched enough to connect with V.

Technical Proof. A dynamical system is said to be *topologically mixing* if, for every two non-empty open sets U and V in X, there exists an N sufficiently large so that $f^n(U) \cap V \neq \varnothing$ for all $n \geq N$.

A measure preserving transformation $T : (X, \Sigma, \mu) \to (X, \Sigma, \mu)$ is *mixing* (with respect to the measure μ) if for any two measurable $A, B \subset X$, we have

$$\mu(T^{-n}(A) \cap B) \to \mu(A) \cdot \mu(B) \text{ as } n \to \infty.$$

It follows that if a measure preserving transformation is mixing, then it is topologically mixing (so long as the open subsets have positive measure). In this case, Lebesgue measure is such that non-empty open sets have positive measure (take a point in the set, then you can find an open interval contained in the open set, and the open interval has positive measure). Hence, for $U, V \subset \mathbb{T}^2$ open and non-empty, if T is mixing with respect to the Lebesgue measure, then $\mu(T^{-n}(A) \cap B) > 0$ for n large enough, implying that $T^{-n}(A) \cap B \neq \varnothing$ for n large enough.

Let H be a Hilbert space. A family of functions $\mathcal{F} \subset H$ is said to be *complete* if there is no function $f \in H$ so that $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{F}$. In other words, $\langle f, \varphi \rangle = 0$ for all $\varphi \in \mathcal{F}$ implies $f = 0$.

We have the following proposition from Katok.

Proposition 2. Consider a measure space (X, Σ, μ). T is mixing (with respect to the measure) if and only if for any complete system $\mathcal{F} \subset L^2(X, \mu)$ of functions and any $f, g \in \mathcal{F}$, we have

$$\int f(T^n(x))\overline{g}(x)d\mu \to \left(\int f d\mu\right) \cdot \left(\int \overline{g} d\mu\right) \text{ as } n \to \infty. \tag{3}$$

Proof. **Step 1:** We show that if condition (1) holds for a complete system $\mathcal{F}$, then it holds for linear combinations of elements in $\mathcal{F}$.

To see this, note that the inner product in L^2 is given by

$$\langle f, g \rangle = \int f \overline{g} d\mu.$$

So we rewrite condition (1) as

$$\langle f \circ T^n, g \rangle \to \left(\int f d\mu\right) (\overline{g} d\mu). \tag{4}$$

Now, if condition (2) holds for all of the complete system, then it holds for linear combinations, since

$$\langle (\alpha_1 f_1 + \alpha_2 f_2) \circ T^n, (\beta_1 g_1 + \beta_2 g_2) \rangle = \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_i \overline{\beta_j} \langle f_i \circ T^n, g_j \rangle,$$

and so

$$\begin{aligned}
\lim_{n\to\infty} \langle (\alpha_1 f_1 + \alpha_2 f_2) \circ T^n, (\beta_1 g_1 + \beta_2 g_2) \rangle &= \lim_{n\to\infty} \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_i \overline{\beta_j} \langle f_i \circ T^n, g_j \rangle \\
&= \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_i \overline{\beta_j} \lim_{n\to\infty} \langle f_i \circ T^n, g_j \rangle \\
&= \sum_{i,j=1}^{2} \alpha_i \overline{\beta_j} \left(\int f_i d\mu\right) \cdot \left(\int \overline{g_j} d\mu\right) \\
&= \left(\int \alpha_1 f_1 + \alpha f_2 d\mu\right) \cdot \left(\int \overline{\beta_1 g_1 + \beta_2 g_2} d\mu\right)
\end{aligned}$$

using the linearity of integrals.

Step 2: It holds for linear combinations of elements in $\mathcal{F}$, so in particular it holds for a dense subset $L(\mathcal{F}) \subset L^2(X,\mu)$. This follows, since the span of $\mathcal{F}$ will contain an orthonormal basis, so will be dense (with respect to the topology generated by the norm $\|f\| = \sqrt{\langle f,f\rangle}$).

Step 3: We now wish to show condition (1) or (2) holds for all $f,g \in L^2(X,\mu)$. Since $L(\mathcal{F})$ is dense, for $\epsilon > 0$ and $f,g \in L^2(x,\mu)$ we can find $f',g' \in L(\mathcal{F})$ so that $\|f-f'\| < \epsilon$, $\|g-g'\| < \epsilon$. We see that (using the Schwarz inequality)

$$\begin{aligned}
&|\langle f\circ T^n, g\rangle - \langle f,g\rangle| \\
&= \Bigg| \langle f\circ T^n, g-g'\rangle + \langle (f-f')\circ T^n, g\rangle + \langle f'\circ T^n, g'\rangle \\
&- \left(\int f' d\mu\right)\left(\int \overline{g'} d\mu\right) + \left(\int f' d\mu\right)\left(\int \overline{g'-g} d\mu\right) + \left(\int (f'-f) d\mu\right)\left(\int \overline{g} d\mu\right)\Bigg| \\
&\leq \|f\circ T^n\|\|g-g'\| + \|(f-f')\circ T^n\|\|g'\| \\
&+ \left|\int f'(T^n(x))\overline{g'}(x) d\mu - \left(\int f' d\mu\right)\left(\int \overline{g'} d\mu\right)\right| + \left|\int f' d\mu\right| \cdot \|g-g'\| + \|f-f'\| \cdot \left|\int \overline{g} d\mu\right| \\
&\leq \left|\int f'(T^n(x))\overline{g'}(x) d\mu - \left(\int f' d\mu\right)\left(\int \overline{g'} d\mu\right)\right| + \epsilon\left(\|f\| + 1 + \left|\int f' d\mu\right| + \left|\int \overline{g} d\mu\right|\right).
\end{aligned}$$

If we take $\epsilon \to 0$, we get the desired result, which is that condition (2) holds.

Step 4: We finish the proof. If T is mixing, then we note that the set $\mathcal{F} = \{\chi_A : A \text{ is measurable}\}$ forms a complete system, so by **Step 3** we have that condition (1) holds for all $L^2(X,\mu)$. If condition (1) holds on a complete system, then by **Step 3** we see that it holds on all of $L^2(X,\mu)$, so in particular it holds on $f = \chi_A$ and $g = \chi_B$ for $A, B \subset X$ measurable, and so T is mixing. $\square$

We now show that T is topologically mixing by showing that T is mixing with respect to Lebesgue measure, which means showing that condition (1) holds on a complete system. The system $\mathcal{F} = \{\exp(2\pi i(m,n)\cdot(x,y)) : (m,n) \in \mathbb{Z}^2\}$ forms a complete system (by Stone-Weierstrass), so we just need to show condition (1) holds on this system. Let

$$\chi_{(m,n)}(x,y) = \exp(2\pi i(m,n)\cdot(x,y)).$$

So we just need to show that

$$\int \chi_{(m,n)}(T^l(x))\overline{\chi_{(j,k)}} d\mu \to \left(\int \chi_{(m,n)} d\mu\right)\cdot\left(\int \chi_{(j,k)} d\mu\right).$$

If $m = n = j = k = 0$, we get that $\chi_{(0,0)} = 1$, so both the left and right sides are 1. Assume $(m,n) \neq (0,0)$. We see that the right hand side of the equation is 0, so it suffices to show that the left hand side converges to 0. Notice that

$$\chi_{(m,n)}(T(x)) = \exp(2\pi i(m,n)\cdot(2x+y, x+y)).$$

We can correspond to the action of T on $\mathbb{T}^2$ the matrix

$$\widehat{T} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

so that

$$\chi_{(m,n)}(T^N(x,y)) = \exp(2\pi i(m,n)\cdot T^N(x,y)) = \exp(2\pi i \cdot \widehat{T}^N(m,n)\cdot(x,y)) = \chi_{\widehat{T}^N(m,n)}(x,y)$$

We can now rewrite the left hand side as

$$\int \chi_{\widehat{T}^N(m,n)-(j,k)} d\mu.$$

Now, this is zero so long as $\widehat{T}^N(m,n) - (j,k) \neq 0$, or $\widehat{T}^N(m,n) \neq (j,k)$. Since $\|\widehat{T}^N(m,n)\| \to \infty$, there exists sufficiently large N so that $\widehat{T}^N(m,n) \neq (j,k)$. So we have that the integral converges to 0. $\square$

Problem 15 (Problem 1.10, James)**.** Let (x_n) be a subadditive sequence of non-negative real numbers. That is,

$$x_{n+m} \leq x_n + x_m.$$

Show that

$$\lim_{n\to\infty} \frac{x_n}{n} = \inf_{n\geq 0} \frac{x_n}{n}.$$

Proof. Let

$$R = \inf_{n\geq 0} \frac{x_n}{n}.$$

For $\epsilon > 0$, there exists an $n \geq 0$ so that

$$\frac{x_n}{n} < R + \epsilon \Leftrightarrow x_n < n(R+\epsilon)$$

For $m \geq n$, write $m = qn + r$, $0 \leq r < n$. Notice that

$$x_m = x_{qn+r} \leq x_{qn} + x_r,$$

and

$$x_{qn} = x_{n+(q-1)n} \leq x_n + x_{(q-1)n},$$

so that recursively applying gives us

$$x_{qn} \leq qx_n.$$

Hence,

$$x_m \leq qx_n + x_r.$$

Let

$$T = \max\{x_i : 0 \leq i < n\}.$$

This then says that

$$x_m \leq qx_n + T.$$

Thus,

$$\frac{x_m}{m} \leq \frac{qx_n + T}{m} = q\frac{x_n}{m} + \frac{T}{m} < \frac{qn}{m}(R+\epsilon) + \frac{T}{m}.$$

So taking the limit, we have

$$\limsup_{m\to\infty} \frac{x_m}{m} \leq R + \epsilon,$$

noting that

$$m = qn + r \Leftrightarrow 1 = \frac{qn}{m} + \frac{r}{m} \Leftrightarrow 1 = \lim_{m\to\infty} \frac{qn}{m}.$$

This holds for all $\epsilon > 0$, so we can take $\epsilon \to 0$ to get

$$\limsup_{m\to\infty} \frac{x_m}{m} \leq \inf_{n\geq 0} \frac{x_n}{n}.$$

Now we have that

$$\inf_{n\geq 0} \frac{x_n}{n} \leq \liminf_{n\to\infty} \frac{x_n}{n} \leq \limsup_{n\to\infty} \frac{x_n}{n} \leq \inf_{n\geq 0} \frac{x_n}{n},$$

so

$$\lim_{n\to\infty} \frac{x_n}{n} = \inf_{n\geq 0} \frac{x_n}{n}.$$

$\square$

Problem 16 (Problem 1.11, James)**.** Show that if $S : Y \to Y$ is a factor of $T : X \to X$, then $h(T) \geq h(S)$.

Proof. Recall that S is a factor of T if we have that there is a surjective continuous map $h : X \to Y$ with

$$h \circ T = S \circ h.$$

Assuming we're on (compact) metric spaces (Y, d_Y) and (X, d_X), we recall

$$h(T) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log(S_d(T, \epsilon, n)),$$

where $S_d(T, \epsilon, n)$ is the cardinality of the minimal (n, ϵ)-spanning set. The idea is to use this definition to show the inequality.

Since h is continuous (notice uniformly continuous since we're on a compact metric space), for all $\epsilon > 0$ there is a $\delta(\epsilon) > 0$ so that

$$d_X(x, y) < \delta(\epsilon) \implies d_Y(h(x), h(y)) < \epsilon.$$

In other words,

$$h(B^X_{\delta(\epsilon)}(x)) \subset B^Y_\epsilon(h(x)).$$

Let $E \subset X$ be an (n, ϵ)-spanning set.

Recall we define

$$d^T_{X,n}(x, y) = \max\{d_X(T^i(x), T^i(y)) : 0 \leq i \leq n\}.$$

By the observation above, for $\epsilon > 0$ fixed and for each iterate $0 \leq i \leq n$, there exists a $\delta(\epsilon, i) > 0$ with

$$d_X(T^i(x), T^i(y)) < \delta(\epsilon, i) \implies d_Y(h(T^i(x)), h(T^i(y))) < \epsilon.$$

Notice

$$h \circ T^i = h \circ T \circ T^{i-1} = S \circ h \circ T^{i-1},$$

so iterating we get

$$h \circ T^i = S^i \circ h.$$

In other words,

$$d_X(T^i(x), T^i(y)) < \delta(\epsilon, i) \implies d_Y(S^i(h(x)), S^i(h(y))) < \epsilon.$$

It's a finite collection, so taking $\delta(\epsilon) = \min\{\delta(\epsilon, i) : 0 \leq i \leq n\}$ gives us that for fixed $\epsilon > 0$ and for all $0 \leq i \leq n$,

$$d_X(T^i(x), T^i(y)) < \delta(\epsilon) \implies d_Y(S^i(h(x)), S^i(h(y))) < \epsilon.$$

Translating this to be in terms of the Bowen-Dinaburg metric,

$$d^T_{X,n}(x, y) < \delta(\epsilon) \implies d^S_{Y,n}(h(x), h(y)) < \epsilon.$$

Taking balls with respect to the Bown-Dinaburg metrics, we have

$$h(B^X_{\delta(\epsilon)}(x)) \subset B^Y_\epsilon(h(x)).$$

Now consider an $(n, \delta(\epsilon))$-spanning set for X. The above observation and surjectivity says taking the image of this gives us a (n, ϵ)-spanning set for Y. So the minimal cardinality for a $(n, \delta(\epsilon))$-spanning set for X will be at least the minimal cardinality for a (n, ϵ)-spanning set for Y. In terms of the S_d, we have

$$S_{d_Y}(S, \epsilon, n) \leq S_{d_X}(T, \delta(\epsilon), n).$$

The monotonicity of logarithms and the independence of n tells us

$$\limsup_{n \to \infty} \frac{1}{n} \log(S_{d_Y}(S, \epsilon, n)) \leq \limsup_{n \to \infty} \frac{1}{n} \log(S_{d_X}(T, \delta(\epsilon), n)).$$

Taking $\epsilon \to 0$ gives

$$h(S) \leq \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log(S_{d_X}(T, \delta(\epsilon), n)) = h(T),$$

as desired. $\square$

Remark. See **Proposition 3.1.6**.

Problem 17 (Problem 1.13, James, Suxuan)**.** Calculate the topological entropy of $T : \mathbb{T}^2 \to \mathbb{T}^2$ defined by

$$T(x, y) = (x, x + y) \pmod 1.$$

Proof. Examining the lines

$$\Lambda_x = \{(x, y) : y \in S^1\}.$$

Note that

$$T(\Lambda_x) \subset \Lambda_x,$$

since T on each of these vertical lines acts just like a rotation. Notice as well that the speed of the rotation depends on the location of the circle on the torus (when viewing $\mathbb{T}^2$ as a quotient of $[0, 1] \times [0, 1]$). As x goes further right, the circle rotates faster under T. However, on the line T is just a rotation, so a continuous isometry.

We will use this interpretation of things to determined an upper bound on $S_d(T, \epsilon, n)$, the minimal number of point so that

$$X \subset \bigcup B_T(x, \epsilon, n).$$

We will do so by cleverly choosing candidate points based on how T is going to act. Vertically, we uniformly choose points distance ϵ apart, since T acts as an isometry. Notice that the d_n^T distance between two points lie along the same horziontal line which are d distance ϵ apart increases by a factor of ϵ for every rotation. So uniformly spread points horizontally across by a d distance of ϵ/n. Then the d_n^T balls of radius ϵ cover the torus. The number of points which covers the torus in this fashion will be n/ϵ^2, so $S_d(T, \epsilon, n) \leq n/\epsilon^2$. Taking the logarithm of this, we have

$$\frac{1}{n} \log(S_d(T, \epsilon, n)) \leq \frac{\log(n) - 2\log(\epsilon)}{n}.$$

Taking the limsup as $n \to \infty$, limit as $\epsilon \to 0$ of both sides gives

$$h(T) = 0.$$

□

Remark.

- If you can show the fact that
$$X \subset \bigcup_x \Lambda_x \implies h(T) = \sup_x h(T_x),$$
you'll get the same calculation (see **Exercise 8.2.5** in **A First Course in Dynamics With a Panorama of Recent Developments** by **Katok & Hasselblatt**.) I could prove such a relation for *countable* unions, but didn't see how to extend it for uncountable unions, and couldn't find a reference either way.
- The inspiration for the solution comes from **Section 8.2.3** in **A First Course in Dynamics With a Panorama of Recent Developments** by **Katok & Hasselblatt** though the details they give are lacking and I think there's a typo.

Problem 18 (Problem 1.14, Hao)**.** Let $f : S^1 \to S^1$ be a degree 2 map. Show that $h(f) \geq \log(2)$.

Proof. We first need a lemma.

Lemma. Let $x \in S^1$. Define m_x as $m_x = y - x$, where $x < y$, $f(x) = f(y)$, and y is the smallest value for which

$$f|_{[x,y]} : [x, y]/(x \sim y) \cong S^1 \to S^1$$

is of degree 1. Then

$$\inf_{x \in S^1} m_x > 0.$$

Proof of Lemma. We proceed by contradiction. Suppose $\inf_{x\in S^1} m_x \leq 0$. Since S^1 compact, we can find a sequence $(x_n) \subset S^1$ with $x_n \to x$ and with $m_{x_n} = y_n - x_n \to 0$ as $n \to \infty$. Taking a subsequence, we may assume $x_n \searrow x$. Let $y > x$ be the smallest value for which $f|_{[x,y]}$ is of degree 1. By definition of our sequence, there is an $n_1 \in \mathbb{N}$ with

$$x < x_{n_1} < y_{n_1} < y,$$

and an $n_2 \in \mathbb{N}$ with

$$x < x_{n_2} < y_{n_2} < x_{n_1} < y_{n_1} < y.$$

Proceeding in this manner, we can find a subsequence (x_{n_k}) with

$$[x_{n_k}, y_{n_k}] \cap [x_{n_j}, y_{n_j}] = \varnothing.$$

Notice that for all $k \in \mathbb{N}$, we have

$$f([x_{n_k}, y_{n_k}]) = S^1.$$

So for every $k \in \mathbb{N}$, we can find $a_k, b_k \in [x_{n_k}, y_{n_k}]$ with $f(a_k) = 0$ and $f(b_k) = 1/2$. Since $y_{n_k} \to x$ as $k \to \infty$, we have that $a_k, b_k \to x$ as well. On the other hand, the sequence $\{f(a_1), f(b_1), f(a_2), f(b_2), \ldots\}$ is a sequence that does not converge. This contradicts the continuity of f. $\square$

As before, this $\inf_{x\in S^1} m_x$ gives us a nice way of bounding the distance between preimages. The goal is to then use (n, ϵ) separated sets to prove our result. Let $0 < \epsilon < \inf_{x\in S^1} m_x$. The goal is to show

$$N_d(f, \epsilon, n) \geq \frac{2^n \pi}{\epsilon}.$$

Going by induction, we have that for the base case we can simply take $2\pi/\epsilon$ points on S^1 which are evenly spaced distance ϵ apart. So

$$N_d(f, \epsilon, 1) \geq \frac{2\pi}{\epsilon}.$$

Now, for the induction step, we use the fact that f has degree 2. Let S be an (n, ϵ)-separated set of size $2^n\pi/\epsilon$. For each $y \in S$, we can find $x_1^y, x_2^y \in S^1$ with $x_1^y < x_2^y$ which satisfies

- $f(x_1^y) = f(x_2^y) = y$,
- $f|_{[x_1^y, x_2^y]} : S^1 \to S^1$ has degree 1.

Let $E \subset S^1$ be given by

$$E = \{x_i^y : i = 1, 2, y \in S\}.$$

Notice that given $y_1, y_2 \in S$ distinct, we have

$$\begin{aligned} d_{n+1}^f(x_i^{y_1}, x_j^{y_2}) = \max_{0\leq k\leq n} d(f^k(x_i^{y_1}), f^k(x_j^{y_2})) &\geq \max_{1\leq k\leq n} d(f^k(x_i^{y_1}), f^k(x_j^{y_2})) \\ &= \max_{0\leq k\leq n-1} d(f^k(y_1), f^k(y_2)) \geq \epsilon. \end{aligned}$$

For $y \in S$, we see that

$$d_{n+1}^f(x_1^y, x_2^y) \geq d(x_1^y, x_2^y) \geq \epsilon.$$

So E is an $(n+1, \epsilon)$-separated set, giving us

$$N_d(f, \epsilon, n+1) \geq \frac{2^{n+1}\pi}{\epsilon}.$$

$\square$

Problem 19 (Problem 1.15, Suxuan)**.** Let $f : \mathbb{T}^2 \to \mathbb{T}^2$ be a map and let $A = f_*$ be the induced matrix on the first homology (or the fundamental group). Prove that $h_{top}(f) \geq \log \lambda$, where λ is the spectral radius of A i.e., the biggest absolute value of the eigenvalues of A.

Proof. Firse we consider the case where $f(x,y) = (ax+by, cx+dz) \pmod 1$ with $a,b,c,d \in \mathbb{Z}$, then $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We assume that A is diagonalizable. Now let us calculate $P_n(f)$. Let $f^n(x,y) = (a_n x + b_n y, c_n x + d_n y) \pmod 1$. If $f^n(x,y) = (x,y)$, then $(a_n - 1)x + b_n y$ and $c_n x + (d_n - 1)y$ are integers, and the map $f_n - id : (x,y) \mapsto ((a_n - 1)x + b_n y, c_n x + (d_n - 1)y)$ is a noninvertible map of the torus onto itself. Let λ_1 and λ_2 be the eigenvalues of A and let D be the orthogonal matrix such that DAD^{-1} is diagonal. Then the number of preimages of $(0,0)$ is

$$|\det(f^n - id)| = |\det((DAD^{-1})^n - id)| = |(\lambda_1^n - 1)(\lambda_2^n - 1)|.$$

□

Problem 20 (Problem 1.16, Hao)**.** Give an example of a degree 2 map $f : S^2 \to S^2$ with zero topological entropy.

Remark. A slight modification from here, **page 149**. Credit to Hao for finding it.

Proof. Identify S^2 with the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Define the map

$$f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$$

by

$$f(z) = \begin{cases} 2\frac{z^2}{|z|} \text{ if } z \neq 0 \\ 0 \text{ if } z = 0. \end{cases}$$

on $\widehat{C}$, we see that f pushes all of the points towards ∞. Since 0 and ∞ are fixed, and otherwise all open neighborhoods are eventually pushed far away from a point, we see $\mathrm{NW}(f) = \{0, \infty\}$. This is a finite set, so $h(f) = h(f|_{\mathrm{NW(f)}}) = 0$. It follows this map has degree 2. □

3. One Dimensional Dynamics

Problem 21 (Problem 2.1, James, Suxuan)**.** Consider the tent map

$$f : [0,1] \to [0,1], \qquad f(x) = 1 - \mid 1 - 2x \mid.$$

Prove that for any $n \geq 1$, there exists a periodic point $p \in [0,1]$ whose smallest period is n.

Remark. This proof is long and technical. Suxuan gave a great geometric proof of this, which is just draw what the tent map looks like at each iteration and then draw the line $f(x) = x$, and note that points of intersection give points whose period is at most n. Deduce that one of them must have period n.

Proof. We design an algorithm which will find a periodic point of period n.

We can write

$$f(x) = \begin{cases} 2x \text{ if } 0 \leq x < \frac{1}{2} \\ 2 - 2x \text{ if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Consider the dyadic rational endpoints,

$$S_k^i = \{\frac{i}{2^k}, 0 \leq i \leq 2^k\}.$$

Then we see that if we iterate f, we have

$$f^2(x) = \begin{cases} 4x \text{ if } 0 \leq x < \frac{1}{4} \\ 2 - 4x \text{ if } \frac{1}{4} \leq x < \frac{1}{2} \\ 4 - 4x \text{ if } \frac{1}{2} \leq x < \frac{3}{4} \\ 4x - 2 \text{ if } \frac{3}{4} \leq x \leq 1. \end{cases}$$

or

$$f^2(x) = \begin{cases} 4x \text{ if } S_2^0 \leq x < S_2^1 \\ 2-4x \text{ if } S_2^1 \leq x < S_2^2 \\ 4-4x \text{ if } S_2^2 \leq x < S_2^3 \\ 4x-2 \text{ if } S_2^3 \leq x \leq S_2^4. \end{cases}$$

Let $g(x) = 2x$ and $h(x) = 2-2x$. We can then rewrite this as

$$f(x) = \begin{cases} g(x) \text{ if } S_1^0 \leq x < S_2^1 \\ h(x) \text{ if } S_1^1 \leq x \leq S_2^2 \end{cases}$$

$$f^2(x) = \begin{cases} g^2(x) \text{ if } S_2^0 \leq x < S_2^1 \\ h(g(x)) \text{ if } S_2^1 \leq x < S_2^2 \\ h^2(x) \text{ if } S_2^2 \leq x < S_2^3 \\ g(h(x)) \text{ if } S_2^3 \leq x \leq S_2^4. \end{cases}$$

We now move on to the case $n = 3$ to get

$$f^3(x) = \begin{cases} g^3(x) \text{ if } S_3^0 \leq x < S_3^1 \\ h(g^2(x)) \text{ if } S_3^1 \leq x < S_3^2 \\ h^2(g(x)) \text{ if } S_3^2 \leq x < S_3^3 \\ g(h(g(x))) \text{ if } S_3^3 \leq x < S_3^4 \\ g(h^2(x)) \text{ if } S_3^4 \leq x < S_3^5 \\ h^3(x) \text{ if } S_3^5 \leq x < S_3^6 \\ h(g(h(x))) \text{ if } S_3^6 \leq x < S_3^7 \\ g^2(h(x)) \text{ if } S_3^7 \leq x \leq S_3^8 \end{cases}$$

We see a $g-h-h-g$ pattern. The argument is a simple induction argument, noticing that we have a tent map and so we decompose it in the corresponding way. The above discussion tells us that

$$P_n(f) \leq 2^n,$$

where $P_n(f) = |\{x \in [0,1] : f^n(x) = x\}|$.

Note that

$$g^n(x) = 2^n x, \qquad h^n(x) = \frac{2+(-2)^{n+1}}{3} + (-2)^n x.$$

We can use the above to calculate the elements of period at most 2:

$$\left\{0, \frac{2}{5}, \frac{4}{5}, \frac{2}{3}\right\}$$

and the elements with period at most 3:

$$\left\{0, \frac{2}{7}, \frac{4}{7}, \frac{6}{7}, \frac{2}{9}, \frac{4}{9}, \frac{6}{9}, \right\}$$

and the elements with period at most 4:

$$\left\{0, \frac{2}{15}, \frac{4}{15}, \frac{6}{15}, \frac{8}{15}, \frac{10}{15}, \frac{12}{15}, \frac{14}{15}, \frac{2}{17}, \ldots \frac{16}{17}\right\}.$$

We notice that there is a pattern of two common denominators; for period 3, we have 7 and 9, for period 2 we have 5 and 3, for period 4 we have 15 and 17.

We claim that the collection of points

$$\left\{\frac{2k}{2^n-1} : 0 \le k \le 2^{n-1}-1\right\} \cup \left\{\frac{2k}{2^n+1} : 0 \le k \le 2^{n-1}\right\}$$

are the points so that $f^n(x) = x$. Notice that this will establish that $P_n(x) = 2^n$.

Taking this set, we order it in increasing order, and then we claim that the ith point x in this set is the unique point in $[S_n^{i-1}, S_n^i)$ which satisfies $f^n(x) = x$. We go by induction, noting it holds in the first 3 steps. So we assume that it holds for n, and we wish to show it holds for $n+1$. Let κ_k be the combinations of gs and hs which agree with f^n on $[S_n^{k-1}, S_n^k]$. If k is odd, then we have that

$$\kappa_k\left(\frac{2\left(\frac{k-1}{2}\right)}{2^n-1}\right) = \kappa_k\left(\frac{k-1}{2^n-1}\right) = \frac{k-1}{2^n-1},$$

and if k is even then we have

$$\kappa_k\left(\frac{k}{2^n+1}\right) = \frac{k}{2^n+1}.$$

Notice that κ_k is going to be a line with slope $\pm 2^n$, where for k even we have slope -2^n and for k odd we have slope 2^n. We can determine what the line is based on this information. We have that for k even

$$\kappa_k(x) = -2^n x - k,$$

and for k odd we have

$$\kappa_k(x) = 2^n x + k - 1.$$

If k is odd, we have that $g(\kappa_k) = f^{n+1}$ on the interval $[S_{n+1}^{2k-2}, S_{n+1}^{2k-1}]$ and $h(\kappa_k) = f^{n+1}$ on the interval $[S_{n+1}^{2k-1}, S_{n+1}^{2k}]$. If k is even, we have $h(\kappa_k) = f^{n+1}$ on $[S_{n+1}^{2k-2}, S_{n+1}^{2k-1}]$ and $g(\kappa_k) = f^{n+1}$ on $[S_{n+1}^{2k-1}, S_{n+1}^{2k}]$. If k is even, we claim that the fixed point for $h(\kappa_k)$ is

$$h\left(\kappa_k\left(\frac{2\left(\frac{2k-2}{2}\right)}{2^{n+1}-1}\right)\right) = h\left(\kappa_k\left(\frac{2k-2}{2^{n+1}-1}\right)\right).$$

Notice that

$$\kappa_k\left(\frac{2k-2}{2^{n+1}-1}\right) = \frac{2^{n+1}-k}{2^{n+1}-1},$$

and so

$$h\left(\kappa_k\left(\frac{2k-2}{2^{n+1}-1}\right)\right) = h\left(\frac{2^{n+1}-k}{2^{n+1}-1}\right) = \frac{2k-2}{2^{n+1}-1}$$

as desired.

The remaining cases follow a similar pattern, and so we omit them.

This gives us $P_n(f) = 2^n$. Notice that the number of points with period n is at least $P_n(f) - \sum_{0\le i<n} P_i(f)$. We calculate

$$\sum_{0\le i<n} P_i(f) = \sum_{i=0}^{n-1} 2^i = 2^n - 1,$$

so there is at least one point with period n for every $n \ge 1$. $\square$

We need two lemmas before moving on.

Lemma (Preimage Lemma)**.** Suppose I and J are closed intervals and $J \subset f(I)$. Then there exists a closed subinterval $K \subset I$ with $J = f(K)$.

Proof. Write $J = [b_1, b_2]$. Since $J \subset f(I)$, there exists c_1, c_2 with $f(c_1) = b_1$, $f(c_2) = b_2$. Without loss of generality, assume $c_1 < c_2$ (the other direction is the same, it just flips the interval). We can define

$$x_1 = \sup\{x \in [c_1, c_2] : f(x) = b_1\}.$$

Then $f(x_1) = b_1$ by continuity, and if $x \in [c_1, c_2]$ is such that $x > x_1$, then $f(x) > f(x_1)$. Define

$$x_2 = \inf\{x \in [x_1, c_2] : f(x) = b_2\}.$$

Then $f(x_2) = b_2$ by continuity, and if $x \in [x_1, a_2]$ is such that $x < x_2$, then $f(x) < b_2$. So $K = [x_1, x_2] \subset I$ is an interval with $f(K) = J$. □

Remark.

- Adapted from **Lemma 7.3** here.
- Note if the containment $J \subset f(I)$ is strict, we can make $K \subset I$ strict.

Lemma (Fixed Point Lemma)**.** Suppose I and J are closed intervals such that $I \subset J$. If $J \subset f(I)$, then f has a fixed point in I.

Proof. By the Preimage Lemma, we see that there is a $K \subset I$ with $f(K) = I$. Write $K = [c_1, c_2] \subset I = [a_1, a_2]$. Note that $a_1 < c_1 < c_2 < a_2$. We have $f(c_1) = a_1$, so $f(c_1) < c_1$, $f(c_2) = a_2$ so $f(c_2) > c_2$, so taking $f(x) - x$ we see there must be a $x_0 \in J \subset I$ with $f(x_0) - x_0 = 0$, i.e. there must be a fixed point in I. □

Problem 22 (Problem 2.2, James)**.** Prove that if $f : [0, 1] \to [0, 1]$ has a point of period 4, then it also has a point of period 2.

Proof. Let x_0 be the point of period 4. We have

$$\mathcal{O}_f(x_0) = \{x_0, f(x_0), f^2(x_0), f^3(x_0)\}.$$

Notice that there must be c, d with

$$f(d) \leq c < d \leq f(c).$$

Without loss of generality, we can assume x_0 is the smallest value in $\mathcal{O}_f(x_0)$ (otherwise relabel so x_0 is the smallest). We have a few cases to consider.

Case 1: $f(x_0) < f^2(x_0) < f^3(x_0)$: Let $c = f^2(x_0)$, $d = f^3(x_0)$, then $f(d) = x_0$ and $f(c) = f^3(x_0)$.
Case 2: $f(x_0) < f^3(x_0) < f^2(x_0)$: Let $c = f(x_0)$, $d = f^3(x_0)$, then $f(c) = f^2(x_0)$ and $f(d) = x_0$.
Case 3: $f^2(x_0) < f(x_0) < f^3(x_0)$: Let $c = f^2(x_0)$, $d = f(x_0)$, then $f(d) = f^2(x_0)$, $f(c) = f^3(x_0)$.
Case 4: $f^2(x_0) < f^3(x_0) < f(x_0)$: Let $c = x_0$, $d = f(x_0)$, then $f(c) = f(x_0)$ and $f(d) = f^2(x_0)$.
Case 5: $f^3(x_0) < f(x_0) < f^2(x_0)$: Let $c = f(x_0)$, $d = f^2(x_0)$, then $f(c) = f^2(x_0)$ and $f(d) = f^3(x_0)$.
Case 6: $f^3(x_0) < f^2(x_0) < f(x_0)$: Let $c = x_0$, $d = f^2(x_0)$, then $f(c) = f(x_0)$ and $f(d) = f^3(x_0)$.

This enumerates all cases, so we see that we must have this property.

Let $w = \min\{c \leq x \leq d : f(x) = x\}$. Notice the interval $[c, w] \neq \varnothing$, since $f(c) \neq c$ so $w > c$. Let $v \in [c, w]$ be such that $f(v) = d$. Notice there must be such a v, since $w = f(w) \leq d \leq f(c)$ and f is continuous. Then $f^2(v) = f(d) \leq c \leq v$. If f has no fixed points in $[a, c]$, then in particular it doesn't fix points of $[a, v]$. Since $f^2(a) \geq a$, the intermediate value theorem says that there exists a point with period 2 in this interval.

If f does have a fixed point in $[a, c]$, let $t = \max\{a \leq x \leq c : f(x) = x\}$. Since $t < c < v$ (where the first inequality is strict since $f(c) \neq c$), we have that the interval $(t, v]$ is non-trivial. Notice that f doesn't fix any points in $(t, v]$ by construction. Let u be a point in $[t, c]$ with $f(u) = c$ (the existence of such a u follows since $f(t) = t < c$, $f(c) \geq d > c$, so intermediate value theorem applies). Then $f^2(u) = f(c) \geq d > u$, and since $f^2(v) \leq v$ we have $f^2(y) = y$ for some point in

$[u,v]$. Because f doesn't fix any points in $[u,v]$, we get that f admits a point with period at least 2. $\square$

Remark. Adapted from **Lemma 2** here.

Adaption of Hao's Proof. Let $a,b,c,d \in \mathcal{O}_f(x_0)$ be fixed, $a \mapsto b \mapsto c \mapsto d \mapsto a$. We show just one case (the rest of the cases are the same and there's a lot to do).

Assume $a < b < c < d$. Consider $I_1 = [a,b]$, $I_2 = [b,c]$, $I_3 = [c,d]$. We have $I_2 \subset f(I_1)$, $I_3 \subset f(I_2)$, $I_1 \cup I_2 \cup I_3 \subset f(I_3)$. Since $I_3 \subset f(I_2)$, we get by the **Preimage Lemma** that there is a $K_2 \subset I_2$ with $f(K_2) = I_3$. Now $I_2 \subset f(I_3)$, so $K_2 \subset I_2 \subset f^2(K_2)$. By the **Fixed Point Lemma**, there exists a fixed point in K_2 with respect to f^2. Notice $f(K_2) \cap K_2 \subset \{c\}$, and this is a point of period 4, so it is impossible for this point to have period 1. We've thus found a point of period 2. $\square$

Problem 23 (Problem 2.3, James)**.** Suppose a continuous map $f : [0,1] \to [0,1]$ has a periodic point of (smallest) period 3. Prove that for any $n \geq 1$ there exists a periodic point $p \in [0,1]$ whose smallest period is n.

Proof. Assume $x_0 \in [0,1]$ is a point with period 3. Then we have

$$\mathcal{O}_f(x_0) = \{x_0, f(x_0), f^2(x_0)\}.$$

Let $a,b,c \in \mathcal{O}_f(x_0)$ be distinct so that $a < b < c$. Then we can write $I_1 = [a,b]$, $I_2 = [b,c]$. Without loss of generality, assume that under f we have $a \mapsto b \mapsto c \mapsto a$. The other cases are similar.

Since $b \mapsto c$, $a \mapsto b$, we have $I_2 \subset f(I_1)$ (using here that intervals map to intervals). Note that $I_1 \cup I_2 \subset f(I_2)$, since $f(b) = c$, $f(c) = a$, so $[a,c] \subset f(I_2)$. Now, $I_2 \subset f(I_2)$, so by the **Fixed Point Lemma** we get there is a fixed point in I_2. Since $I_2 \subset f(I_1)$, the **Preimage Lemma** says there is a $K_1 \subset I_1$ with $f(K_1) = I_2$. Since $I_1 \subset f(I_2)$, this implies $K_1 \subset I_1 \subset f^2(K_1)$, so the **Fixed Point Lemma** says there is a fixed point for f^2 in K_1. Note that this cannot have period 1, since the only point that K_1 and $f(K_1)$ could possibly share is $\{b\}$, which has period 4.

Note that if we have a point with period 2, we get for free a point of period 1 by the intermediate value theorem. So it suffices to show that for $n > 3$, we can find a point with (minimal) period n. The idea is to iterate the strategy we've just done.

We have $I_2 \subset f(I_2)$, so the **Preimage Lemma** says that there is a $K_1 \subset I_2$ with $f(K_1) = I_2$. Since $K_1 \subset f(K_1)$, we can find a closed interval $K_2 \subset K_1$ with $f(K_2) = K_1$, so $f^2(K_2) = I_2$. Continue in this fashion up to $n-2$, so $K_{n-2} \subset K_{n-1}$ is a closed interval with $f^{n-2}(K_{n-2}) = I_2$. Since $K_{n-2} \subset I_2 \subset f(I_1)$, there is a closed interval $K_{n-1} \subset I_1$ with $f(K_{n-1}) = K_{n-2}$. Since $K_{n-1} \subset I_1 \subset f(I_2)$, there is a closed interval $K_n \subset I_2$ with $f(K_n) = K_{n-1}$. Since $K_n \subset f^n(K_n)$, there is a fixed point for f^n in K_n. Let x be this point.

Note that $f(x) \in I_1$ while $x, f^k(x) \in I_2$ for $1 < k \leq n$. Clearly f cannot have period 1, since the only point shared between I_1 and I_2 has period 3. For the other cases, we utilize the remark after the **Preimage Lemma** which says that we can make these sets strictly decreasing. $\square$

Remark. See Sharkovskii's Theorem.

Problem 24 (Problem 2.4, James)**.** Let $f : S^1 \to S^1$ be the "times 2" map $x \mapsto 2x \pmod 1$. Let $\epsilon > 0$ be small and (x_n) be a sequence of points such that

$$d(x_{n+1}, f(x_n)) < \epsilon.$$

(1) Prove that there exists an orbit $f^n(z)$ such that

$$d(f^n(z), x_n) < 2\epsilon.$$

(2) Prove that such a z is unique.

We need two claims.

Claim. For $n \geq 1$, if $d(x,y) < 2^{-(n+1)}$, then $d(f(x), f(y)) < 2^{-n}$.

Proof. If $d(x,y) < 2^{-(n+1)}$, then we have

$$|x-y| < 2^{-(n+1)} \text{ or } 1 - |x-y| < 2^{-(n+1)}.$$

If $x, y < 1/2$, then $f(x) = 2x$, $f(y) = 2y$, and $d(x,y) = |x-y| < 2^{-(n+1)}$. So

$$|f(x) - f(y)| = 2|x-y| < 2 \cdot 2^{-(n+1)} = 2^{-n} \implies d(f(x), f(y)) < 2^{-n}.$$

If $x, y \geq 1/2$, then $f(x) = 2x-1$, $f(y) = 2y-1$, and $d(x,y) = |x-y| < 2^{-(n+1)}$. So

$$|f(x) - f(y)| = 2|x-y| < 2^{-n} \implies d(f(x), f(y)) < 2^{-n}.$$

If $x < 1/2$, $y \geq 1/2$, then $f(x) = 2x$, $f(y) = 2y-1$. Notice

$$d(f(x), f(y)) \leq 1 - |f(x) - f(y)| = 1 - |2x - (2y-1)|,$$
$$d(f(x), f(y)) \leq |f(x) - f(y)| = |2x - (2y-1)|.$$

If $d(x,y) = |x-y|$, then this implies that $x \in [1/4, 1/2)$, $y \in [1/2, 3/4]$. So $f(x) \in [1/2, 1)$ and $f(y) \in [0, 1/2]$. In other words, $f(x) \geq f(y)$, so $|f(x) - f(y)| = f(x) - f(y)$. Hence,

$$d(f(x), f(y)) \leq 1 - (2x - (2y-1)) = 2x - 2y \leq 2|x-y| < 2^{-n}.$$

If $d(x,y) = 1 - |x-y|$, then this implies $x \in [0, 2^{-(n+1)})$ and $y \in (1 - 2^{-(n+1)}, 1)$, so $f(x) \in [0, 2^{-n}]$ and $f(y) \in (1 - 2^{-n}, 1)$, forcing $d(f(x), f(y)) < 2^{-n}$. $\square$

Remark.

- As a corollary, we have that if $\epsilon < 1/2$, then

  $$d(x,y) < \epsilon \implies d(f(x), f(y)) < 2\epsilon.$$

  One can refine this to say that for sufficiently close points, the distance between them doubles for iterates of f.
- As a reference, the contents of this claim can be found in the proof of **Katok & Hasselblatt, Proposition 3.2.3**.

Claim. For $0 < \delta < 1/16$ we have that

$$d(f(x), y) \leq 2\delta \implies \overline{B_\delta(x)} \cap f^{-1}(y) \neq \varnothing.$$

In other words, there is a z with $d(x,z) \leq \delta$ and $f(z) = y$.

Proof. This follows by using the fact that δ is sufficiently small and we're dealing with an expanding map. By the observation above, this implies

$$f(\overline{B_\delta(x)}) = \overline{B_{2\delta}(f(x))},$$

so $y \in f(\overline{B_\delta(x)})$, meaning there is a $z \in \overline{B_\delta(x)}$ with $f(z) = y$. $\square$

Remark. Though not directly related, the inspiration for the solution came from **Proposition 1** found here.

Solution of Problem. We break it up into steps. Throughout, fix $0 < \epsilon < 1/32$. Note that in particular $2\epsilon < 1/16$.

Step 1: Let

$$A_n := \{z \in S^1 : f^i(z) \in \overline{B_\epsilon(x_i)} \text{ for } 0 \leq i \leq n\}, \qquad A := \bigcap_{n \geq 0} A_n.$$

The idea is to show that each of these A_n are nonempty and closed. This implies that they are compact, and since $A_m \subset A_n$ for $m \geq n$ we get that the FIP implies $A \neq \varnothing$. The layout is as follows:

- Step 2 will show each A_n is nonempty.
- Step 3 will show each A_n is closed. From here, we deduce A is nonempty.
- Step 4 will then show that A is just one point.

Step 2: By assumption, we have

$$d(f(x_i), x_{i+1}) < \epsilon$$

for $0 \leq i$. For $j \geq 1$ set

$$\lambda_1 = 1, \qquad \lambda_j = \frac{\lambda_{j-1}}{2} + 1.$$

Solving this recursion, we get

$$\lambda_j = 2 - 2^{1-j}.$$

By assumption, for fixed n we have

$$d(f(x_{n-1}), x_n) < \epsilon.$$

The goal is to build a sequence $z_j^{(n)}$ where each $z_j^{(n)}$ satisfies our desired property. Set $z_n^{(n)} = x_n$. Using our above claim, there exists a $z_{n-1}^{(n)} \in \overline{B_{\epsilon/2}(x_{n-1})}$ with $f(z_{n-1}^{(n)}) = x_n$. We see that

$$\begin{aligned} d(f(x_{n-2}), z_{n-1}^{(n)}) &\leq d(f(x_{n-2}), x_{n-1}) + d(x_{n-1}, z_{n-1}^{(n)}) \\ &< \epsilon\left(1 + \frac{1}{2}\right) = \lambda_2\epsilon = \frac{3}{2}\epsilon < 2\epsilon. \end{aligned}$$

We now iterate. Notice that there is a $z_{n-2}^{(n)} \in \overline{B_{\lambda_2\epsilon/2}(x_{n-2})}$ with $f(z_{n-2}^{(n)}) = z_{n-1}^{(n)}$ and

$$d(f(x_{n-3}), z_{n-2}^{(n)}) \leq d(f(x_{n-3}), x_{n-2}) + d(x_{n-2}, z_{n-2}^{(n)}) < \epsilon\left(\frac{\lambda_2}{2} + 1\right) = \lambda_3\epsilon = \frac{7}{4}\epsilon < 2\epsilon.$$

We can find a $z_{n-3}^{(n)} \in \overline{B_{\lambda_3\epsilon/2}(x_{n-3})}$ with $f(z_{n-3}^{(n)}) = z_{n-2}^{(n)}$ and

$$d(f(x_{n-4}), z_{n-3}^{(n)}) \leq d(f(x_{n-4}), x_{n-3}) + d(x_{n-3}, z_{n-3}^{(n)}) \leq \epsilon\left(\frac{\lambda_3}{2} + 1\right) = \lambda_4\epsilon = \frac{15}{8}\epsilon < 2\epsilon.$$

Continuing inductively, we see that for $2 \leq i \leq n$, we can find $z_{n-i}^{(n)} \in \overline{B_{\lambda_i\epsilon/2}(x_{n-i})}$ with

$$f(z_{n-i}^{(n)}) = z_{n-i+1}^{(n)} \text{ and } d(f(x_{n-i-1}), z_{n-i}^{(n)}) \leq \lambda_{i+1}\epsilon = (2 - 2^{1-i-1})\epsilon < 2\epsilon.$$

Notice $z_0^{(n)}$ satisfies $f^j(z_0^{(n)}) = z_j^{(n)}$ for $0 \leq j \leq n$ by construction. We see then

$$d(f^{n-j}(z_0^{(n)}), x_{n-j}) = d(z_{n-j}^{(n)}, x_{n-j}) \leq \frac{\lambda_j}{2}\epsilon \leq \epsilon \text{ for } 0 \leq j \leq n.$$

So A_n is nonempty for each n.

Step 3: Fix $n \geq 0$. We now show A_n is closed by taking sequences. Let $(z_m) \subset A_n$ be a sequence of points with $z_m \to z$. The goal is to show $z \in A_n$ as well. Since $z_m \to z$, we get that for all $\kappa > 0$ there exists sufficiently large N with

$$d(z_m, z) < \kappa$$

for $m \geq N$. So

$$d(x_0, z) \leq d(x_0, z_m) + d(z_m, z) < \epsilon + \kappa.$$

This is for all $\kappa > 0$, so taking $\kappa \to 0$ we get

$$d(x_0, z) \leq \epsilon.$$

For $0 \le i \le n$, we see

$$d(f^i(z), x_i) \le d(f^i(z), f^i(z_m)) + d(f^i(z_m), x_i).$$

Since f^i is continuous on the circle, we can take m sufficiently large so that for all $\kappa > 0$

$$d(f^i(z), x_i) \le d(f^i(z), f^i(z_m)) + d(f^i(z_m), x_i) < \kappa + \epsilon.$$

Taking $\kappa \to 0$, we have

$$d(f^i(z), x_i) \le \epsilon.$$

So it holds for all $0 \le i \le n$. This shows that A_n is a closed set, so in particular it is compact. Putting everything together, there exists a $z \in A$ so that

$$d(f^i(z), x_i) \le \epsilon$$

for all i.

Step 4: Suppose $y_0, y_1 \in A$. Then for all n, we have

$$d(f^i(y_0), x_i) \le \epsilon.$$

This coupled with the triangle inequality implies

$$d(f^i(y_0), f^i(y_1)) \le 2\epsilon$$

for all i. Since $\epsilon < 1/32$, this says that

$$d(f^i(y_0), f^i(y_1)) < \frac{1}{4}$$

for all $i \ge 0$. We claim that this is impossible unless $y_0 = y_1$.

If $y_0 \ne y_1$, then we have $d(y_0, y_1) = \delta > 0$. By an earlier observation, the distance between two points doubles until it passes $1/4$. So we have

$$2^i \delta < \frac{1}{4}$$

for all i. But this is impossible, since this implies $\delta = 0$, a contradiction. Hence A is a single point.

So we have that there is a z so that

$$d(f^i(z), x_i) \le \epsilon < 2\epsilon.$$

□

Remark. A slight modification will give you that this property holds for all expanding maps.

4. Ergodic Theory

Problem 25 (Problem 3.1, James)**.** Let $T : (X, \mu) \to (X, \mu)$ be a measure preserving transformation. Prove that the following are equivalent.

(1) T is ergodic.
(2) If f is measurable and $f \circ T = f$, then f is constant almost everywhere.
(3) If f is measurable and $f \circ T = f$ almost everywhere, then f is constant almost everywhere.
(4) If $f \in L^2(X, \mu)$ and $f \circ T = f$, then f is constant almost everywhere.
(5) If $f \in L^2(X, \mu)$ and $f \circ T = f$ almost everywhere, then f is constant almost everywhere.

Remark. Recall that $T : (X, \mu) \to (X, \mu)$ is *ergodic* if and only if for any measurable set $A \subset X$ with $T^{-1}(A) = A$, we have either $\mu(A) = 0$ or $\mu(A^c) = 0$.

Proof. (1) $\implies$ (2): Let $f : X \to \mathbb{R}$ be a measurable function, $T : X \to X$ ergodic, and suppose $f \circ T = f$ on all of X. For $c \in \mathbb{R}$, consider the sets $A_c = \{x \in X : f(x) \geq c\}$ and $B_c = \{x \in X : f(x) < c\}$. Since $f \circ T = f$, we have

$$A_c = T^{-1}(A_c) = \{x \in X : f \circ T(x) \geq c\}.$$

By ergodicity, this implies $\mu(A_c) = 0$ or $\mu(B_c) = 0$. If f were not constant almost everywhere, then we have that there is some $c \in \mathbb{R}$ with $\mu(A_c),\ \mu(B_c) \neq 0$, contradicting the above property. Hence f is constant almost everywhere.
(2) $\implies$ (3): Suppose that $f : X \to \mathbb{R}$ is measurable and $f \circ T = f$ almost everywhere, say on a set $A \subset X$. We can define a new function $g : X \to \mathbb{R}$ such that $g = f$ almost everywhere and $g \circ T = g$ everywhere. We have that g is a measurable function (assuming the measure is complete) being almost everywhere equal to a measurable function, and by (2) we see that g is constant almost everywhere, so this forces f to be constant almost everywhere.
(3) $\implies$ (4): Suppose $f \in L^2(X, \mu)$ with $f \circ T = f$ everywhere. In particular f is measurable, and we see that $f \circ T = f$ almost everywhere, so f is constant almost everywhere.
(4) $\implies$ (5): The same trick for (2) $\implies$ (3) applies here.
(3) $\implies$ (1): Let $A \subset X$ be a measurable set with $T^{-1}(A) \subset A$. We have that $\chi_A : X \to \mathbb{R}$ is a measurable function and we note that $\chi_A \circ T = \chi_{T^{-1}(A)} = \chi_A$. Hence property (3) says that χ_A is constant almost everywhere, implying that either $\mu(A) = 0$ or $\mu(A^c) = 0$.

I believe (3) $\iff$ (4) $\iff$ (5) depends on whether the space has finite measure, but I cannot find a reference for this. $\square$

Problem 26 (Problem 3.2, James)**.** Let $R : x \mapsto x + \alpha$ be the rotation map on the circle.

(1) Prove that R is ergodic.
(2) Prove that R is uniquely ergodic.

Proof.

(1) First, we show that a rotation is measure preserving. Note that the measurable sets w.r.t. Lebesgue measure are generated by open intervals, so it suffices to show that for any interval $(a, b) \subset [0, 1]$, we have $m(R^{-1}((a, b))) = m((a, b))$. We note that

$$\begin{aligned} R^{-1}((a,b)) &= \{x \in [0,1] : R(x) = x + \alpha \pmod 1 \in (a,b)\} \\ &= (a - \alpha, b - \alpha) \pmod 1. \end{aligned}$$

If $b - \alpha < 0$, then we get that this is mapped to $(a - \alpha + 1, b - \alpha + 1)$, and the measure of this is $b - a$. If there exists a $c \in (a, b)$ with $c - \alpha = 0$, then this is mapped to $(a - \alpha + 1, c - \alpha + 1] \sqcup [c - \alpha, b - \alpha)$, which has measure $b - a$ again. Finally, if $a - \alpha \geq 0$, then this is mapped to $(a-\alpha, b-\alpha)$, which has measure $b-a$. Hence it is measure preserving on the generators for the σ-algebra, so measure preserving.

Now we need to show ergodicity. We use **Problem 2.1**. This is a finite measure space so the equivalence of all the properties holds. In particular, let $f : X \to \mathbb{R}$ be in $L^2(S^1, m)$, and suppose that $f \circ R = f$ almost everywhere. Since f is in L^2, we can write it in terms of its Fourier series;

$$f(x) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n x},$$

$$f \circ R(x) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n (x + \alpha)} = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n x} e^{2\pi i n \alpha},$$

where a_n are the Fourier coefficients. Since these are equal almost everywhere, their Fourier coefficients are equal, so

$$a_n e^{2\pi i n \alpha} = a_n \Rightarrow a_n \left[1 - e^{2\pi i n \alpha}\right] = 0.$$

Thus either $a_n = 0$ or $e^{2\pi i n\alpha} = 1$. The latter only happens if $n\alpha$ is an integer. Since α is irrational, this only happens if $n = 0$. Thus, we have that $f(x) = a_0$ almost everywhere, so f is constant almost everywhere. This tells us that R is ergodic.

(2) A map is *uniquely ergodic* iff it has one ergodic Borel probability measure. Let m denote Lebesgue measure. Let μ be a measure such that R is ergodic with respect to μ. The goal is to show that for any $f \in C(\mathbb{T})$ we have

$$\int f dm = \int f d\mu.$$

Riesz-Representation then says that $m = \mu$ as measures, and we can conclude uniqueness.

Let

$$a_n = \int f(x) e^{-2\pi i n x} dm$$

be the Fourier coefficients with respect to Lebesgue measure. For $f \in C(X)$, we denote the partial sums as

$$\sigma_N(x) = \sum_{-N}^{N} a_n e^{2\pi i n x}.$$

Notice that we have

$$\int \sigma_N(x) d\mu = \int \left(\sum_{-N}^{N} a_n e^{2\pi i n x} \right) d\mu = a_0 = \int f dm.$$

To see this explicitly, we have

$$\int e^{2\pi i k x} d\mu = \int e^{2\pi i k (x+\alpha)} d\mu = e^{2\pi i k \alpha} \int e^{2\pi i k x} d\mu$$

using the fact that μ is T-invariant. Since α is irrational we see that this integral must be zero for $k \neq 0$. Proceeding by linearity gives the above result.

Using Fourier, we see that $\sigma_N \to f$ uniformly, so that

$$\int f dm = \lim_{n\to\infty} \int \sigma_n d\mu = \int f d\mu.$$

This gives us the desired result.

$\square$

Problem 27 (Problem 3.3, James)**.** Give an example of a uniquely ergodic homeomorphism of a compact metric space which is not minimal.

Proof. See this blog post. $\square$

Problem 28 (Problem 3.4, James, Suxuan)**.** Prove that $T : \mathbb{T}^2 \to \mathbb{T}^2$ given by $T(x, y) = (x + \alpha, y + x)$ for $\alpha \notin \mathbb{Q}$ is ergodic.

Proof. We first need to check that T preserves measure with respect to Lebesgue measure. We can equivalently show that, for all $f : \mathbb{T}^2 \to \mathbb{R}$ which are integrable, we have

$$\int f(x, y) d(x \times y) = \int f \circ T d(x \times y).$$

Notice

$$\int f \circ T(x, y) d(x \times y) = \int f(x + \alpha, x + y) d(x \times y).$$

Fubini/Tonelli applies to give

$$\iint_{\mathbb{T}^2} f(x + \alpha, x + y) dy dx.$$

Let $z = x + y$, $dz = dy$, we see that

$$\iint_{\mathbb{T}^2} f(x+\alpha, x+y)\,dy\,dx = \iint_{\mathbb{T}^2} f(x+\alpha, z)\,dz\,dx,$$

and doing a change of variables $u = x + \alpha$, we have

$$\iint_{\mathbb{T}^2} f(u,z)\,dz\,du = \int_{\mathbb{T}^2} f(x,y)\,d(x \times y).$$

Thus, it is measure preserving.

Next, we need to establish ergodicity. Let $f \in L^2(\mathbb{T}^2, \lambda)$ be such that $f \circ T = f$ almost everywhere. The goal is to establish that f is constant almost everywhere. We again use Fourier series. On the torus, we note that

$$f(x,y) = \sum_{n \in \mathbb{Z}^2} a_n e^{2\pi i n \cdot (x,y)} = \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i n_1 x} e^{2\pi i n_2 y}$$

almost everywhere. Note that

$$\begin{aligned} f \circ T(x,y) = f(x+\alpha, x+y) &= \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i n_1 (x+\alpha)} e^{2\pi i n_2 (x+y)} \\ &= \sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1,n_2)} e^{2\pi i (n_1+n_2) x} e^{2\pi i n_1 \alpha} e^{2\pi i n_2 y}. \end{aligned}$$

Doing a shift in n_1 (since it ranges over all integers anyways), we rewrite the series to get

$$\sum_{n_1,n_2=-\infty}^{\infty} a_{(n_1-n_2,n_2)} e^{2\pi i n_1 x} e^{2\pi i (n_1-n_2)\alpha} e^{2\pi i n_2 y}.$$

The same trick as before applies. The coefficients of these series must be equal, so we have

$$a_{(n_1,n_2)} = a_{(n_1-n_2,n_2)} e^{2\pi i (n_1-n_2)\alpha}.$$

Notice that we have $|a_{(n_1,n_2)}| = |a_{(n_1-n_2,n_2)}| = |a_{(n_1-2n_2,n_2)}| = \cdots = |a_{(n_1-kn_2,n_2)}| = \cdots$, where k is an integer. Riemann-Lebesgue forces $a_{(n_1,n_2)} = 0$ if $n_2 \neq 0$. Thus, it suffices to examine the case $n_2 = 0$. Here, we have

$$a_{(n_1,0)} = a_{(n_1,0)} e^{2\pi i n_1 \alpha}.$$

Notice that this is the same as

$$a_{(n_1,0)} \left[1 - e^{2\pi i n_1 \alpha}\right] = 0,$$

so either $a_{(n_1,0)} = 0$ or $e^{2\pi i n_1 \alpha} = 1$. Like before, α is irrational, so this implies that for no non-zero integer n_1 we have $e^{2\pi i n_1 \alpha} = 1$, so $a_{(n_1,0)} = 0$. Hence, the function is constant almost everywhere. $\square$

Suxuan's Proof. For $f \in L^2(\mathbb{T}^2)$, since $m(\mathbb{T}^2) = 1 < \infty$, by Hölder's inequality, we have

$$\|f\|_1 \leq \|m(\mathbb{T}^2)^{\frac{1}{2}}\|f\|_2 < \infty,$$

so $f \in L^1(\mathbb{T}^2)$. Let $f(x,y) = \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i (mx+ny)}$ be the Fourier series of f, then

$$\begin{aligned} f(T(x,y)) &= \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i (m(x+\alpha)+n(y+x))} \\ &= \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i m\alpha} e^{2\pi i ((m+n)x+ny)}. \end{aligned}$$

Suppose $f(x,y) = f(T(x,y))$, then we have

$$\sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i(mx+ny)} = \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i m\alpha} e^{2\pi i((m+n)x+ny)},$$

so the coefficients of the term $e^{2\pi i(kx+ly)}$, $(k,l)\in\mathbb{Z}$ are the same, so

$$a_{(m+n,n)} = e^{2\pi i m\alpha} a_{(m,n)}.$$

Set $n=0$, then

$$a_{(m,0))} = e^{2\pi i m\alpha} a_{(m,0)},$$

and since $\alpha\notin\mathbb{Q}$, $e^{2\pi i m\alpha}\neq 0$ for $m\in\mathbb{Z}\setminus\{0\}$, so $a_{(m,0)}=0$ for nonzero integers m.
For $n\neq 0$, we have

$$a_{(m,n)} = e^{-2\pi i m\alpha} a_{(m+n,n)} = \cdots = e^{-2\pi i m r\alpha} a_{(m+rn,n)} = \cdots.$$

By the Riemann-Lebesgue lemma, $a_{(m+rn,n)}\to 0$ as $r\to\infty$, and since $|e^{-2\pi i m r\alpha}|\leq 1$, $e^{-2\pi i m r\alpha} a_{(m+rn,n)}\to 0$ as $r\to\infty$, so $a_{(m,n)}=0$ for $n\neq 0$. Then we obtain $f(x,y)=a_{(0,0)}$ a.e. Hence T is ergodic. □

Problem 29 (Problem 3.5, James, Suxuan)**.** Prove that $T:\mathbb{T}^2\to\mathbb{T}^2$ given by $T(x,y)=(2x+y,x+y)$ is ergodic.

Proof. It suffices to check that for $f:\mathbb{T}^2\to\mathbb{R}$ integrable, we have

$$\int f(x,y)d(x\times y) = \int f\circ T(x,y)d(x\times y).$$

Lebesgue measure is σ-finite, so it is fine to use the lemma. Note that by definition of T, we have

$$\int_{\mathbb{T}^2} f\circ T(x,y)d(x\times y) = \int_{\mathbb{T}^2} f(2x+y,x+y)d(x\times y).$$

Note that Fubini-Tonelli works here, so it is fine to iterate the integral to get

$$\iint_{\mathbb{T}^2} f(2x+y,x+y)dxdy.$$

A change of variables $u=x+y$, $z=x+u$ gives $du=dy$, $dz=dx$, and so we get the integral is equal to

$$\iint_{\mathbb{T}^2} f(z,u)dzdu = \int_{\mathbb{T}^2} f(x,y)d(x\times y).$$

Hence, T is measure preserving.

We now need to show that T is ergodic with respect to Lebesgue measure. We invoke the L^2 definition of ergodic, so T is ergodic iff for all $f\in L^2(\mathbb{T}^2,\lambda)$, we have that $f\circ T=f$ implies f is constant almost everywhere. We have that f in $L^2(\mathbb{T}^2,\lambda)$ implies that we can express it almost everywhere in terms of its Fourier series,

$$f(x,y) = \sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)} e^{2\pi i(n_1,n_2)\cdot(x,y)}$$

so

$$\begin{aligned} f\circ T(x,y) = \sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)} e^{2\pi i(n_1,n_2)\cdot(2x+y,x+y)} &= \sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)} e^{2\pi i n_1(2x+y)} e^{2\pi i n_2(x+y)} \\ &= \sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1,n_2)} e^{2\pi i x(2n_1+n_2)} e^{2\pi i y(n_1+n_2)}. \end{aligned}$$

Let $k = n_1 + n_2$, then $n_2 = k - n_1$, so the series can be expressed as

$$\sum_{(n_1,k)\in\mathbb{Z}^2} a_{(n_1,k-n_1)} e^{2\pi i x(n_1+k)} e^{2\pi i y k}.$$

Letting $u = n_1 + k$, we have $n_1 = u - k$, so

$$\sum_{(u,k)\in\mathbb{Z}^2} a_{(u-k,2k-u)} e^{2\pi i x u} e^{2\pi i y k}.$$

Relabeling u and k, we have that the series is equivalent to

$$\sum_{(n_1,n_2)\in\mathbb{Z}^2} a_{(n_1-n_2,2n_2-n_1)} e^{2\pi i x n_1} e^{2\pi i y n_2},$$

so since these are equal (almost everywhere), we have that the Fourier series agree, so

$$a_{(n_1,n_2)} = a_{(n_1-n_2,2n_2-n_1)}.$$

Fixing $(n_1, n_2) \in \mathbb{Z}^2$ non-trivial, we iterate to see that

$$a_{(n_1,n_2)} = a_{(n_1-n_2,2n_2-n_1)} = a_{(2n_1-3n_2,5n_2-3n_1)} = \cdots.$$

We see its equal to infinitely many distinct coefficients, so Riemann-Lebesgue says that it must be 0. Hence, it's constant almost everywhere, so T is ergodic. □

Suxuan's Proof. We use the same setting as problem 3.4. $f(x, y) = \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i(mx+ny)}$ be the Fourier series of f, and

$$\begin{aligned} f(T(x,y)) &= \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i(m(2x+y)+n(x+y))} \\ &= \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i((2m+n)x+(m+n)y)}. \end{aligned}$$

Suppose $f(x, y) = f(T(x, y))$, then

$$\sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i(mx+ny)} = \sum_{(m,n)\in\mathbb{Z}^2} a_{(m,n)} e^{2\pi i((2m+n)x+(m+n)y)},$$

so $a_{(2m+n,m+n)} = a_{(m,n)}$. The eigenvalues of the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is $\frac{3+\sqrt{5}}{2} > 1$ and $\frac{3-\sqrt{5}}{2} < 1$, so for $(m, n) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, (m, n) is not an eigenvector of the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Let $\begin{pmatrix} p_k \\ q_k \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^k \begin{pmatrix} m \\ n \end{pmatrix}$, then $|p_k| + |q_k| \to \infty$ as $k \to$. Then by the Riemann-Lebesgue lemma, $a_{(p_k,q_k)} \to 0$ as $k \to \infty$. Since $a_{(m,n)} = a_{(p_k,q_k)}$, we then obtain $a_{(m,n)} = 0$ for $(m, n) \neq (0, 0)$, therefore $f(x, y) = a_{(0,0)}$ a.e. It follows that T is ergodic. □

Problem 30 (Problem 3.6, James)**.** Let $A \in \mathrm{SL}(d, \mathbb{Z})$ and let $A : \mathbb{T}^d \to \mathbb{T}^d$ be the induced automorphisms. Prove that A is ergodic if and only if A has no roots of unity among its eigenvalues.

Proof. Note that A will always be measure preserving, since $\det(A) = 1$ (use one of the equivalent definitions for measure preserving).
$(\Longleftarrow)$: Assume for all $\lambda \in \mathrm{Spec}(A)$, $|\lambda| \neq 1$. We wish to show that A is ergodic. Let $f \in L^2(\mathbb{T}^d)$ be such that $f \circ A = f$ almost everywhere. Using Fourier series, we see that

$$f = \sum_{n\in\mathbb{Z}^d} a_n e^{2\pi i n\cdot x} = \sum_{n\in\mathbb{Z}^d} a_n e^{2\pi i n\cdot Ax} = \sum_{n\in\mathbb{Z}^d} a_n e^{2\pi i n A^t\cdot x}.$$

By a shift, we get that the Fourier coefficients are such that.

$$a_n = a_{A^t n}.$$

A and A^t share the same eigenvalues. The significance of the fact that for all $\lambda \in \mathrm{Spec}(A), |\lambda| \neq 1$ is that the orbit of any non-zero vector $n \in \mathbb{Z}^d$ is either going to be infinite under A. In other words, we get infinitely many coefficients, all equal to each other, or we get that it contracts to 0, meaning all the coefficients are equal to 0. In the first case, Riemann-Lebesgue says that these must all be zero. Combining the cases, we have that only the zero vector is non-trivial, telling us that f is constant almost everywhere.
($\implies$): We go by contrapositive. Assume that A has a root of unity among its eigenvalues; say it is an nth root of unity. This implies that there is a vector $v \in \mathbb{R}^d$ with $A^n v = v$. The map $A^n - \mathrm{Id} : \mathbb{R}^d \to \mathbb{R}^d$ is going to have non-trivial kernel then, and so similarly the transpose of the induced map on $\mathbb{Z}^d$ will have a non-trivial kernel $(A^n)^t - \mathrm{Id} : \mathbb{Z}^d \to \mathbb{Z}^d$, so there is some $0 \neq m \in \mathbb{Z}^d$ with $(A^n)^t m = m$. Hence, it cannot be ergodic; we can construct a function f so that $f \circ A = f$ almost everywhere but f is not constant. □

Problem 31 (Katok & Hasselblatt 4.1.5, James)**.** Suppose $f : X \to X$ is a topologically transitive continuous map of a compact space X and for every continuous function φ the averages

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

converge uniformly; i.e. for all $\varphi \in C(X)$

$$A_n(\varphi) \xrightarrow{\|\cdot\|_\infty} A(\varphi).$$

Prove that f is uniquely ergodic.

Proof. Since f is topologically transitive, there exists x with $\mathcal{O}_f(x)$ dense in X. By definition, we have uniform convergence

$$A_n(\varphi) \to A(\varphi) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi \circ f^k.$$

The goal is to show

$$A(\varphi) = \int \varphi d\mu$$

is constant. Doing so implies that for any other invariant ergodic measure ν,

$$\int \varphi d\nu = \int \varphi d\mu \text{ for all } \varphi \in C(X) \implies \mu = \nu.$$

So μ is uniquely ergodic. To do so, we show that for all other $y \in X$, $A(\varphi)(y) = A(\varphi)(x)$. We claim that for all $y \in \mathcal{O}_f(x)$, we have $A(\varphi)(y) = A(\varphi)(x)$. To see this, notice $y \in \mathcal{O}_f(x)$ implies $y = f^j(x)$ for some j. Hence

$$A(\varphi)(y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(y)) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^j(f^k(x))) = \int \varphi \circ f^j d\mu_x = \int \varphi d\mu_x = A(\varphi)(x),$$

using f-invariance of $d\mu_x$. Hence $A(\varphi)$ is constant on a dense subset $\mathcal{O}_f(x)$. By uniform convergence we have that $A(\varphi)$ is constant everywhere, giving us that the measure is unique. □

Problem 32 (Problem 4.1, James)**.** Let

$$\Sigma_n = \{1, 2, \dots, n\}^{\mathbb{Z}}$$

be the space of sequences and let

$$d((x_n), (y_n)) = 2^{-\min\{|i| : x_i \neq y_i\}}.$$

Prove that d is a metric and that it generates the product topology.

Proof. The first step is to show that this is a metric. There are four properties to show:

(1) $d(x_n, y_n) \geq 0$ by construction.
(2) If $d(x_n, y_n) = 0$, then this implies that for each n, $x_n = y_n$, since otherwise we have an n so that $x_n \neq y_n$, and so
$$d(x_n, y_n) \geq 2^{-|n|} > 0,$$
a contradiction. Likewise, if $x_n = y_n$, then $d(x_n, y_n) = 0$.
(3) Notice
$$d(x_n, y_n) = 2^{-\min\{|i| : x_i \neq y_i\}} = 2^{-\min\{|i| : y_i \neq x_i\}} = d(y_n, x_n).$$
(4) Let $(x_n), (y_n), (z_n)$ be three sequences. Then we claim that
$$d(x_n, y_n) \leq d(x_n, z_n) + d(z_n, y_n).$$
By definition,
$$d(x_n, z_n) = 2^{-\min\{|i| : x_i \neq z_i\}},$$
$$d(y_n, z_n) = 2^{-\min\{|i| : y_i \neq z_i\}},$$
$$d(x_n, z_n) + d(y_n, z_n) = 2^{-\min\{|i| : x_i \neq z_i\}} + 2^{-\min\{|i| : y_i \neq z_i\}}.$$
Let $a = \min\{|i| : x_i \neq y_i\}$, $b = \min\{|i| : x_i \neq z_i\}$, $c = \min\{|i| : y_i \neq z_i\}$. If $b \leq a$ or $c \leq a$, then
$$2^b \leq 2^a \implies 2^{-a} \leq 2^{-b}$$
or
$$2^{-a} \leq 2^{-c},$$
so
$$2^{-a} \leq 2^{-c} + 2^{-b}.$$
Now, we claim that it is impossible for $b > a$ and $c > a$. If $b > a$ and $c > a$, then this says that the smallest integer (in magnitude) where (z_n) disagrees with (x_n) is b, which is farther out than a, and the smallest integer where (z_n) disagrees with (y_n) is c, which is farther out than a. But this means that at index a, z_n agrees with both x_n and y_n, which is impossible since x_n doesn't agree with y_n. So the triangle inequality holds.

Now we need to show that it generates the product topology. On $\mathbb{Z}/N\mathbb{Z}$, we just take the discrete topology, and so we equip Ω_N with the product topology from this. Consider the open ball

$$B_\epsilon(x_n) = \{y_n : d(x_n, y_n) < \epsilon\}.$$

Notice that for $\epsilon > 0$ there exists an n large enough so that $2^{-n} < \epsilon$. Let n be the smallest such integer. Then this says that $d(x_n, y_n) \leq 2^{-n}$, so that this is the collection of all sequences which agree with x_n up to the nth index. So in other words, the open balls are just

$$B_N(x_n) = \{(y_n) : y_j = x_j \text{ for } |j| < N\}.$$

The basic open sets with respect to the product topology are generated by the cylinders of rank k;

$$C^{m_1, \dots, m_k}_{a_1, \dots, a_k} = \{(y_n) : y_{m_j} = a_j\}.$$

Let τ be the product topology, τ_d be the topology generated by d. We see that every open ball corresponds to a cylinder of rank $2N+1$, so $\tau_d \subset \tau$. It suffices to show that we can construct arbitrary cylinders from open balls.

For any m, we need to show that we can construct

$$C_a^m = \{(y_n) : y_m = a\}$$

from open balls. Finite intersections will give us arbitrary cylinders, completing the proof. We can clearly do this for $m = 0$, so assume $m \neq 0$. Let $R = \{(x_n) \in \Omega_N\}$ be the collection of sequences with $x_m = a$. Then

$$\bigcup_{(x_n) \in R} B_{m+1}(x_n) = \{(y_n) : y_m = a\} = C_a^m.$$

□

Problem 33 (Problem 4.2, James, Hao)**.**

(1) Prove that the full shift on n symbols is topologically mixing.
(2) Give an example of an uncountable transitive subshift of finite type on 2 symbols which is not topologically mixing.

Proof.

(1) We collect some definitions and thoughts first.
- We define
$$\Omega_N := \{\omega = (\omega_i)_{i \in \mathbb{Z}} : \omega_i \in \mathbb{Z}/N\mathbb{Z}\}.$$
- We define a *cylinder of rank k* to be the set
$$C_\alpha^n = C_{\alpha_1, \ldots \alpha_k}^{n_1, \ldots, n_k} := \{\omega \in \Omega_N : \omega_{n_i} = \alpha_i \text{ for } 1 \leq i \leq k\},$$
where $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $n = (n_1, \ldots, n_k)$. The topology on Ω_N is generated by the base
$$\mathcal{B} := \{C_\alpha^n : k \in \mathbb{Z}, n \in \mathbb{Z}^k, \alpha \in (\mathbb{Z}/N\mathbb{Z})^{2k+1}\}.$$
- A *symmetric cylinder of rank 2m+1* is defined by
$$S_\alpha^m := C_{\alpha_{-m}, \alpha_{-m+1}, \ldots, \alpha_m}^{-m, -m+1, \ldots, m}.$$
One thing to remark is that for every cylinder C_α^n, there exist a symmetric cylinder S_β^m so that $S_\beta^m \subset C_\alpha^n$. Hence, we can refine our base to be
$$\mathcal{B} = \{S_\alpha^m : m \in \mathbb{Z}, \alpha \in (\mathbb{Z}/N\mathbb{Z})^{2m+1}\}.$$
- The *left shift* $\sigma_N : \Omega_N \to \Omega_N$ given by $\sigma_N(\omega) = \omega'$, where $\omega'_i = \omega_{i+1}$. As noted in Katok, this is a homeomorphism on Ω_N.
- A 0-1 matrix A is said to be *transitive* if there exists a positive integer m so that the entries of A^m are all greater than 0.
- Let $m = \min\{k : a_{ij}^k > 0 \text{ for all } i, j\}$. We call m the *transitive power of A*.
- If m is the transitive power of A, then we claim that for all $n \geq m$ and all vertices i and j we have $a_{ij}^n > 0$. To prove this, we go by induction. Note that it holds for m by construction, so the base case holds. Assume it holds for some $n - 1 \geq m$, we wish to show it holds for n.
Note that we can interpret the entries a_{ij}^{n-1} in A^{n-1} as indicating whether there is a path of length n connecting vertex i to vertex j. If $a_{ij}^n > 0$ for all vertices i and j, we get that there is a path of length n connecting every vertex to every vertex. The claim then is that for any vertex j, there is a path $i = i_0 \to \cdots \to i_{n+1} = j$ connecting i to j. If not, this means that for any choice of vertex i_n there is no edge connecting i_n to j. But this means that there is no edge connecting any vertex k to vertex j, contradicting

the fact that we have a path connecting i to j of length n. Thus, we have some vertex k so that there is an edge (kj), and we have a path of length $n-1$ connecting i to k, so adding this vertex gives a path of length n connecting i to j. The choice of vertices i and j were arbitrary, so we get that $a_{ij}^n > 1$ for all vertices i, j. (Note that Katok offers an analytic approach instead, see **Lemma 1.9.7**).

- Let $A := (a_{ij})_{i,j=0}^{N-1}$ be a matrix where $a_{ij} \in \{0, 1\}$. We call this a *0-1 matrix.*
- Let
$$\Omega_A := \{\omega \in \Omega_N : a_{\omega_n \omega_{n+1}} = 1 \text{ for } n \in \mathbb{Z}\}.$$
Consider the directed graph $G_A = (V, E)$, where $V = \{0, \dots, N-1\}$ and $(ij) \in E$ iff $a_{ij} = 1$. We note that A is then the transition matrix for G_A. Note that we can then interpret Ω_A to be the space of infinite admissible closed paths on the graph G_A.
- The restriction $\sigma_A := \sigma_N|_{\Omega_A}$ is the *subshift of finite type.*
- A *full n-shift* is σ_A where $A = (a_{ij})$ is such that $a_{ij} = 1$.
- The *symmetric cylinder of rank 2k+1 on* Ω_A is defined to be
$$S_{\alpha,A}^k := \Omega_A \cap S_\alpha^k.$$
Equipping Ω_A with the subspace topology, we have that the base
$$\mathcal{B}' := \{S_{\alpha,A}^k : k \in \mathbb{Z}, \alpha \in (\mathbb{Z}/N\mathbb{Z})^{2k+1}\}$$
generates the topology on Ω_A.
- Recall that a topological dynamical system $f : X \to X$ is said to be *topologically mixing* if, for every U, V open and nonempty, there exists an N so that for all $n > N$ we have $f^n(U) \cap V \neq \varnothing$.
- If the topology of X admits a base $\mathcal{B}_X$, then to show that $f : X \to X$ is topologically mixing it suffices to show that for every $U, V \in \mathcal{B}$, there exists an N so that for all $n > N$ we have $f^n(U) \cap V \neq \varnothing$. This follows since $\mathcal{B}_X$ is a base, so for every nonempty U and V there is $A, B \in \mathcal{B}_X$ so that $A \subset U$, $B \subset V$, and we get the result.
- To show that σ_A is topologically mixing, we can reduce the problem even further. Let $S_{\alpha,A}^k$, $S_{\beta,A}^m$ be two symmetric cylinders. Letting $s = \min\{k, m\}$ we have that we can project α and β to the middle $2s+1$ coordinates. Denote these projections by α', β' respectively. Then we have that $S_{\alpha',A}^s \subset S_{\alpha,A}^k$ and $S_{\beta',A}^s \subset S_{\beta,A}^k$. If we show it for these new cylinders, then we get the result holds for the larger cylinders. Hence it suffices to check that the mapping is topologically mixing on cylinders of the same rank.

Utilizing the above definitions and remarks, it suffices to show that if $S_{\alpha,A}^k$ and $S_{\beta,A}^k$ are two symmetric cylinders, then there exists an N so that for $n > N$ we have
$$\sigma_A^n(S_{\alpha,A}^k) \cap S_{\beta,A}^k \neq \varnothing.$$
To simplify notation, denote
$$S_\alpha := S_{\alpha,A}^k, \quad S_\beta := S_{\beta,A}^k.$$
Since σ_A is a full shift, let $N = 2k+1$ and take $n > N$. We have that n is of the form $n = 2k+1+j$ for $j > 0$ an integer. Since α_k, β_{-k} are in $\mathbb{Z}/N\mathbb{Z}$, we see that $a_{\alpha_k \beta_{-k}}^j > 0$; in other words, there is a closed path of length j connecting the vertex α_k to β_{-k}. Hence, we can construct an admissible path of length $(2k+1)+(2k+1)+j = 4k+2+j$, where the first $2k+1$ coordinates are the vector α and the last $2k+1$ coordinates are the vector β, and the middle j coordinates are the ones which connect α_k to β_{-k}. Setting the rest of the coordinates to whatever we wish, we see that this will be in the intersection so that $\sigma_A^n(S_\alpha) \cap S_\beta \neq \varnothing$. Hence, the full n shift is topologically mixing.

(2) It's on two symbols, so we're looking at

$$\Omega_A = \{\omega \in \Omega_2 : a_{\omega_n \omega_{n+1}} = 1 \text{ for } n \in \mathbb{Z}\}$$

for some matrix A. For the subshift to be transitive, we need the matrix A to also be transitive. By the same sort of argument as in (1), I believe that if the *matrix* is transitive, then we always get σ_A is topologically mixing (see Katok **Proposition 1.9.9**). However, we can construct a *topologically transitive* subshift which is not topologically mixing. Let

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

Then

$$A^n = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } n \text{ is even,} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{if } n \text{ is odd.} \end{cases}$$

so A is not transitive. Ω_A is the set of sequences which alternate 0 and 1. A is irreducible since we can always find a path connecting two vertices. Since every sequence maps to every sequence under σ_A, we get that the system is actually minimal. It is not mixing, since it will not stay in any symmetric cylinder, as it is alternating. Hao has the less trivial counterexample.

TODO: FIND HAOS EXAMPLE

□

6. Smooth Dynamics

Problem 34 (Problem 5.1)**.** Prove that the topological entropy of a diffeomorphism of a compact manifold is finite.

Remark. Result can be found here. The problem says solved on the website but I (James) must have missed this week.

Problem 35 (Problem 5.3, James)**.** Let $f : S^1 \to S^1$ be the map $f(x) = 2x \pmod 1$ and let $g : S^1 \to S^1$ be a C^1 map which commutes with f. Prove that g is a times m map for some m.

Proof. We have $f \circ g = g \circ f$ on S^1. In particular, $2g(0) = g(0)$, so $g(0) = 0$. Next, notice that

$$(f(g(x)))' = f'(g(x))g'(x) = g'(f(x))f'(x) = (g(f(x)))'.$$

Hence, we have

$$2g'(x) = 2g'(2x) \implies g'(x) = g'(2x).$$

Taking $x/2$, we see that this equivalently says

$$g'(x/2) = g'(x).$$

Iterating this, we see

$$g'(x) = g'(2^{-1}x) = \cdots = g'(2^{-n}x)$$

for all n. By the continuity of the derivative, we can take the limit as $n \to \infty$ to get

$$g'(x) = g'(0)$$

for all $x \in S^1$. Hence, we have $g'(x)$ is constant on S^1, which forces $g(x) = mx \pmod 1$ for $m = g'(0)$. □

Remark. This is similar to **Lemma 2.1.4**.

Problem 36 (Problem 5.4, Hao)**.** Let M be a Riemannian manifold. Let $f : M \to M$ be an expanding map. That is, there is a $\lambda > 1$, $C > 0$ such that for all $v \in TM$ and $n \geq 1$, we have

$$\|Df^n v\| \geq C\lambda^n \|v\|.$$

Let $\epsilon > 0$ be small and $(x_i)_{i\geq 0}$ be a sequence of points such that $d(x_{i+1}, f(x_i)) < \epsilon$. Prove that there exists an orbit $f^i(z)$ such that $d(f^i(z), x_i) < C\epsilon$, where C is a constant depending only on λ. Prove that such a z is unique.

We introduce two lemmas first.

Lemma (Lemma 1)**.** Let M be a compact manifold, $f : M \to M$ a local diffeomorphism. Then f is a covering map onto its image.

Proof of Lemma 1. We break this up into parts.

Step 1: We need to show that for all $y \in M$, $f^{-1}(y)$ is finite. To do so, notice that $\{y\}$ is closed in M, so $f^{-1}(y)$ is a closed subset of M. Since M is compact, this implies $f^{-1}(y)$ is compact. For each $x \in f^{-1}(y)$, take a neighborhood U_x of x such that f is a local diffeomorphism on U. Notice that $\{U_x\}_{x\in f^{-1}(y)}$ is an open cover of $f^{-1}(y)$, and so by compactness there are finitely many of these sets; label them U_i for $1 \leq i \leq n$. Restricted to each U_i, f is a diffeomorphism, so in particular injective, and hence U_i consists of exactly one point. Let $x_i \in U_i$, then $\{x_1, \ldots, x_n\} = f^{-1}(y)$.

Step 2: Keep the $\{U_i\}_{i=1}^n$ which are neighborhoods around each x_i. After slightly shrinking them, we can consider these all pairwise disjoint (here we utilize the Hausdorff property). Let

$$V = \bigcap_{i=1}^{n} f(U_i).$$

Then $\{f^{-1}(V) \cap U_i\}$ is a disjoint collection of open neighborhoods homeomorphic to V under f. So V is evenly covered neighborhood of y.

Step 2 tells us that this is a covering map. □

Before proving the next lemma, we recall the Lebesgue number lemma.

Theorem (Lebesgue Number Lemma)**.** If (x, d) is a compact metric space, $\{U_\alpha\}$ is an open cover of X, then there is a number $\delta > 0$ such that every subset of X having diameter less than δ is contained in some member of the cover.

Lemma (Lemma 2)**.** Let M be a compact Riemannian manifold, $f : M \to M$ expanding. Then there exists an $\epsilon > 0$ such that for all $y = f(x) \in M$, there is a $U \subset M$ open and $x \in U$ with

(1) $f(U) = B_\epsilon(y)$,
(2) $f|_U : U \to f(U)$ is a diffeomorphism,
(3) $f|_U$ is strictly expanding.

Proof of Lemma 2. Since the map is expanding, we claim there is an ϵ_0 sufficiently small for which $f|_{B_{\epsilon_0}(x)}$ is strictly expanding for any x (see **Proposition 2.4.2**; this utilizes the Implicit Function Theorem and compactness). Let $y \in f(M)$ (the image of M). Choose a neighborhood V_y of y which satisfies two properties:

(1) $f^{-1}(V_y) = \bigsqcup_{x\in f^{-1}(y)} U_x$, with U_x a neighborhood of x.
(2) For each $x \in f^{-1}(y)$, $\operatorname{diam}(U_x) < \epsilon_0$ so that f is strictly expanding on U_x.

Property (1) is by **Lemma 1**, and property (2) is by the remark earlier. Notice the collection of V_y cover M. We can invoke the Lebesgue number lemma to find an $\epsilon > 0$ such that for any

$y' = f(x') \in f(M)$, there is a $B_\epsilon(y') \subset V_y$ for some y. Using property (1), we see

$$f^{-1}(B_\epsilon(y')) \subset f^{-1}(V_y) = \bigsqcup_{x \in f^{-1}(y)} U_x.$$

Intersecting gives us "slices," $S_x = f^{-1}(B_\epsilon(y')) \cap U_x$. One of these slices must contain x' (where again, $f(x') = y'$). Notice this slice satisfies the properties of the lemma. $\square$

With this, we have enough to prove the problem.

Proof of Problem. Let $\epsilon' > 0$ be given as in Lemma 2. Let $\epsilon < \epsilon'/2$. The goal is to define a sequence (y_i). in the first step, we set $y_0 = x_0$. We move on to the inductive step. Given y_i such that $f^i(y_i) = x_i$, we construct y_{i+1} as follows. Since

$$d(f(x_i), x_{i+1}) = d(f^{i+1}(y_i), x_{i+1}) < \epsilon,$$

there exists y_{i+1} such that $f^{i+1}(y_{i+1}) = x_{i+1}$, and y_{i+1} in the slice of $f^{-i}(B_\epsilon(f^i(y_i))) = f^{-i}(B_\epsilon(x_i))$ containing y_i. To do so, we need to iteratively apply f^{-1}.

Throughout the next paragraph, assume we choose things sufficiently small so that we can ignore any C factor as being problematic.

Lemma 1 says that f is a covering map onto its image. So since ϵ small, we take $f^{-1}(B_\epsilon(x_{i+1}))$ and it divides up into finitely different open sets (we refer to these as "slices"). Since $f^{i+1}(y_i)$ is in $B_\epsilon(x_i)$, there is a slice U containing $f^i(y_i)$. Take $z_i \in U$ with the property that $f(z_i) = x_{i+1}$. Notice that

$$d(z_i, f^i(y_i)) < \frac{1}{\lambda} d(f(z_i), f^{i+1}(y_i)) = \frac{1}{\lambda} d(x_{i+1}, f^{i+1}(y_i)) < \frac{\epsilon}{C\lambda}.$$

Since $\lambda > 1$, this is smaller than ϵ/C (which, taking ϵ small, we can ignore as being problematic). We can now take the ball of radius $\epsilon/C\lambda$ around z_i. We see $f^i(y_i)$ is in this, so there is a slice such that $f^{i-1}(y_i)$ is in that slice. In that slice, we choose z_{i-1} so that $f(z_{i-1}) = z_i$. Notice

$$d(z_{i-1}, f^{i-1}(y_i)) < \frac{1}{\lambda} d(z_i, f^i(y_i)) < \frac{\epsilon}{C\lambda^2}.$$

This tells us that we can keep iterating this process, so we can continue until we get z_0 which satisfies the property that $f^{i+1}(z_0) = x_{i+1}$, and which lives in the same slice as y_i. Label $y_{i+1} = z_0$.

Notice by the locally contracting property, we have

$$d(y_i, y_{i+1}) < \frac{\epsilon}{C\lambda^{i+1}}.$$

We note this implies the sequence is Cauchy, hence converging. To see this, take i, j arbitrary, and notice (assuming wlog $j \geq i$)

$$\begin{aligned} d(y_i, y_j) \leq d(y_j, y_{j-1}) + \cdots + d(y_{i+1}, y_i) &< \frac{\epsilon}{C}\lambda^{-i+1} \sum_{k=0}^{j-i-1} \lambda^{-k} \\ &< \frac{\epsilon}{C}\lambda^{-i+1} \frac{1 - \lambda^{i-j}}{1 - \lambda}. \end{aligned}$$

Taking i, j arbitrarily large makes this arbitrarily small as desired. So there is some $z \in M$ with $y_i \to z$.

The goal now is to show

$$d(f^i(z), f^i(y_i)) = d(f^i(z), x_i) < C'\epsilon,$$

where C' is a constant depending only on λ. To do this, we will show

$$d(f^i(y_j), f^i(y_{j+1})) < \frac{\epsilon}{C\lambda^{j+1-i}}.$$

If we show this, we have $\lim_{j\to\infty} f^i(y_j) = f^i(z)$ implies

$$d(f^i(z), f^i(y_i)) \le \sum_{j=i}^{\infty} d(f^i(y_j), f^i(y_{j+1})) < \frac{\epsilon}{C(1-(1/\lambda))}.$$

Setting $C' = 1/(C(1-1/\lambda))$ gives the desired result.

To see this, notice we have $f^{j+1}(y_j) = f(x_j) \in B_\epsilon(x_{j+1})$, $f^i(y_j) = f^i(y_j) \in B_{\epsilon/(C\lambda^{j+1-i}}(f^i(y_{j+1}))$, and $f^{j+1-i}(f^i(y_j)) = f(x_j) \in B_\epsilon(x_{j+1})$.

Uniqueness is the usual argument; that is,

$$d(f^i(z), f^i(w)) < 2\epsilon,$$

and then use the contracting property to get that

$$d(z,w) < 2\epsilon/(C\lambda^i).$$

Taking $i \to \infty$ gives the result. □

Problem 37 (Problem 5.5, James)**.** Assume that M is a compact Riemannian manifold. Assume $f : M \to M$ is an expanding map with respect to a Riemannian metric $\|\cdot\|$ and let $\|\cdot\|'$ be some other Riemannian metric. Prove that f is also expanding with respect to $\|\cdot\|'$.

Proof. Take $\|\cdot\|, \|\cdot\|'$ to be two Riemannian metrics on M. The goal is to show that they are equivalent in the sense that there are constants $c, C > 0$ so that

$$c\|\cdot\| \le \|\cdot\|' \le C\|\cdot\|.$$

If this can be shown, then since f is expansive with respect to $\|\cdot\|$ we have that there are constants $\lambda > 1$, $K > 0$ so that

$$\|Df^n v\| \ge K\lambda^n\|v\|.$$

Utilizing the above gives us

$$\|Df^n v\|' \ge c\|Df^n v\| \ge cK\lambda^n\|v\| \ge \frac{cK}{C}\lambda^n\|v\|'.$$

So f is expansive with respect to $\|\cdot\|'$ as well. Moreover, the expansive constant λ is the same (though the other constant varies with the metric).

To prove the above claim, we note that since M is compact, UTM is compact (where UTM is the unit tangent bundle). We can then define a function $\gamma : UTM \to \mathbb{R}$ by

$$\gamma(v) = \frac{\|v\|}{\|v\|'}.$$

This functions is continuous and strictly positive (since we're on the unit tangent bundle with respect to the metric $\|\cdot\|$). By compactness, there must be a minimum and a maximum, so constants $c, C > 0$ with

$$c \le \gamma \le C \implies c\|\cdot\|' \le \|\cdot\| \le C\|\cdot\|.$$

□

Remark. Credit to Thomas Richard for the compactness result. See here.

Problem 38 (2.6.1 Katok and Hasselblatt, James)**.** Show that the proof of C^1 structural stability for hyperbolic linear automorphisms can be generalized to any hyperbolic linear automorphism of the m-torus for $m \ge 2$.

Proof. We start by recalling some of the key definitions from this section.

- Recall the definition of structural stability. We say a C^1 map is C^1 structurally stable if there exists a neighborhood U of f in the C^1 topology such that every map $g \in U$ is topologically conjugate to f; that is, for every $g \in U$, there is a homeomorphism h so that $h \circ f = g \circ h$.
- Recall that a hyperbolic linear automorphism is given in the form $F_L : \mathbb{T}^m \to \mathbb{T}^m$, where L denotes the matrix on $\mathbb{R}^m$ and $F_L = T \pmod 1$. A hyperbolic linear automorphism is a linear map on the torus with $L \in \mathrm{GL}_n(\mathbb{R})$, $|\det(L)| = 1$, and for all $\lambda \in \sigma(L)$, $|\lambda| \neq 1$ (that is, the spectra is disjoint from S^1).
- A map on $\mathbb{R}^m$, say g, is doubly periodic if $g(x+m) = g(x)$ for $m \in \mathbb{Z}^m$, $x \in \mathbb{R}^m$.

The goal is to use the contraction mapping principle to find a homeomorphism which conjugates g and F_L, where g is "sufficiently close" to F_L (the closeness to be determined). The equivalent proposition would be **Proposition 2.6.2**, which we reiterate now.

Proposition 3. Any C^1 map g sufficiently close to F_L in the C^1 topology is a factor of F_L.

A lift of F_L is L. A lift of g to $\mathbb{R}^m$ is given by $\widetilde{g}+L$, $\widetilde{g}$ doubly periodic. A lift of h to $\mathbb{R}^m$ is given by $\mathrm{Id}+\widetilde{h}$, $\widetilde{h}$ doubly periodic.

The relation

$$g \circ h = h \circ F_L$$

can be rewritten as

$$\begin{aligned}
(L+\widetilde{g}) \circ (\mathrm{Id}+\widetilde{h}) &= (\mathrm{Id}+\widetilde{h}) \circ L \\
\Leftrightarrow L(\mathrm{Id}+\widetilde{h}) + \widetilde{g}(\mathrm{Id}+\widetilde{h}) &= L + \widetilde{h} \circ L \\
\Leftrightarrow L + L \circ \widetilde{h} + \widetilde{g}(\mathrm{Id}+\widetilde{h}) &= L + \widetilde{h} \circ L \\
\Leftrightarrow \widetilde{g}(\mathrm{Id}+\widetilde{h}) &= \widetilde{h} \circ L - L \circ \widetilde{h}.
\end{aligned}$$

We define two operators on the space of doubly periodic functions on $\mathbb{R}^m$. The first is given by

$$\mathcal{L}(\widetilde{h}) = \widetilde{h} \circ L - L \circ \widetilde{h}.$$

The second is given by

$$\mathcal{T}(\widetilde{h}) = \widetilde{g} \circ (\mathrm{Id}+\widetilde{h}).$$

We can again rewrite the above relation as

$$\mathcal{L}(\widetilde{h}) = \mathcal{T}(\widehat{h}).$$

The first claim is that $\mathcal{L}$ is an invertible operator. To show this, we need to use the expanding/-contracting subspaces of $\mathbb{R}^m$.

For a real eigenvalue λ, we denote by E_λ the subspace

$$E_\lambda = \{v \in \mathbb{R}^m : (A - \lambda \mathrm{Id})^k v = 0 \text{ for some } k\}.$$

For two conjugate complex eigenvalues λ and $\overline{\lambda}$, we denote by $E_{\lambda,\overline{\lambda}}$ the intersection of $\mathbb{R}^m$ with the sum of root spaces corresponding to E_λ, $E_{\overline{\lambda}}$ in the complexification of L. We can then denote by E^- the contracting subspace; i.e. the subspace

$$E^- = E^-(L) = \bigoplus_{|\lambda|<1} E_\lambda \oplus \bigoplus_{|\lambda|<1} E_{\lambda,\overline{\lambda}}.$$

Similarly,

$$E^+ = E^+(L) = \bigoplus_{|\lambda|>1} E_\lambda \oplus \bigoplus_{|\lambda|>1} E_{\lambda,\overline{\lambda}}.$$

Recall for a hyperbolic linear map we have $\mathbb{R}^m = E^+ \oplus E^-$. So every $v \in \mathbb{R}^m$ can be decomposed as $v = v^+ + v^-$, where $v^+ \in E^+$ and $v^- \in E^-$. Denote by λ the eigenvalue which is largest in modulus; that is, $\lambda \in \sigma(L)$ is such that $|\lambda| \geq |\gamma|$ for all $\gamma \in \sigma(L)$. We can express $L = L_1 \times L_2$, where L_i are linear maps with $L_1 : E^+ \to E^+$ and $L_2 : E^- \to E^-$ satisfying the property that for all $v \in \mathbb{R}^m$,

$$L(v) = L(v^+ + v^-) = L(v^+) + L(v^-) = L_1(v^+) + L_2(v^-).$$

If we denote by $\pi_1 : \mathbb{R}^m \to E^+$, $\pi_2 : \mathbb{R}^m \to E^-$, we can write

$$\widetilde{h} = \pi_1(\widetilde{h}) + \pi_2(\widetilde{h}).$$

Write $h_i = \pi_i(\widetilde{h})$, and likewise $g_i = \pi_i(\widetilde{g})$. Then

$$\begin{aligned}\mathcal{L}(\widetilde{h}) = \widetilde{h} \circ L - L \circ \widetilde{h} &= (h_1 + h_2) \circ L - (L_1 + L_2) \circ (h_1 + h_2) \\ &= h_1 \circ L + h_2 \circ L - L_1 \circ h_1 - L_2 \circ h_2.\end{aligned}$$

We can then write new operators

$$\mathcal{L}_1(h_1) = h_1 \circ L - L_1 \circ h_1,$$
$$\mathcal{L}_2(h_2) = h_2 \circ L - L_2 \circ h_2,$$

so that

$$\mathcal{L}(\widetilde{h}) = \mathcal{L}_1(h_1) + \mathcal{L}_2(h_2).$$

The claim is that $\mathcal{L}_i(h_i)$ are invertible for each i. Once we show this, it implies $\mathcal{L}$ is invertible as well. Write

$$\mathcal{L}_1^{-1}(h_1) = -\sum_{n=0}^{\infty} L_1^{-(n+1)} \circ h_1 \circ L^n,$$

then

$$\begin{aligned}\mathcal{L}_1 \circ \mathcal{L}_1^{-1}(h_1) &= \lim_{N\to\infty} \left[-\sum_{n=0}^{N} L_1^{-(n+1)} \circ h_1 \circ L^{n+1} + \sum_{n=0}^{N} L_1^{-n} \circ h_1 \circ L^n \right] \\ &= h_1 - \lim_{N\to\infty} L_1^{-N-1} \circ h_1 \circ L^{N-1} = h_1\end{aligned}$$

using **Proposition 1.2.8**. A similar argument shows $\mathcal{L}_1^{-1} \circ \mathcal{L}_1(h_1) = h_1$, and a similar argument also shows

$$\mathcal{L}_2^{-1}(h_2) = \sum_{n=0}^{\infty} L_2^n \circ h_2 \circ L^{-(n+1)}$$

is true inverse. Since these are invertible, so $\mathcal{L}$ is invertible, we can write rewrite the relation

$$\mathcal{L}(\widetilde{h}) = \mathcal{T}(\widetilde{h}) \Rightarrow \widetilde{h} = \mathcal{L}^{-1} \circ \mathcal{T}(\widetilde{h}).$$

So $\widetilde{h}$ satisfies the criteria if it is a fixed point of the operator $\mathcal{L}^{-1} \circ \mathcal{T}$. Like before, we see

$$\|\mathcal{T}(h) - \mathcal{T}(h')\| = \|\widetilde{g}(\mathrm{Id} + h) - \widetilde{g}(\mathrm{Id} + h')\|_\infty \leq \|D\widetilde{g}\|_\infty \|h - h'\|_\infty,$$

and so

$$\|\mathcal{L}^{-1}\mathcal{T}\| = \|g\|_{C^1} \|\mathcal{L}^{-1}\|,$$

where we note $\|\mathcal{L}^{-1}\|$ depends only on L. So if we have that $\|g\|_{C^1} < \|\mathcal{L}^{-1}\|^{-1}$, then $\mathcal{L}^{-1}\mathcal{T}$ is a contracting operator, and we can apply the contraction mapping principle to get a unique fixed point. Project this fixed point $\mathrm{Id} + \widetilde{h}$ onto the Torus to get a solution. $\square$

References

[1] Anatole Katok & Boris Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems.*
[2] Anatole Katok & Boris Hasselblatt. *A First Course in Dynamics with a Panaroma of Recent Developments.*
[3] Gerald Folland. *Real Analysis : Modern Techniques and Their Applications.* Second Edition.