Exercise 1.1.2

Suppose (X, f) is a factor of (Y, g) by a semi-conjugacy $\pi \colon Y \to X$. (1)

Show that if $y \in Y$ is a periodic point, then $\pi(y) \in X$ is periodic. (2)

Give an example to show that the preimage of a periodic point does not necessarily contain a periodic point. (3)

Proof + reasoning:

Let's start with (1). (4)

Let $y \in Y$ be periodic. (5)

We can just write out the definitions and see if the correct answer follows. (6)

$$f(\pi(y)) \;=\; \pi\big(g(y)\big) \;=\; \pi(y). \tag{7}$$

So, (1) follows from (7). (8)

Now, we want to construct a counterexample. Let's think of any semiconjugacy and iterate from there. An obvious example is a projection. (9)

Let A and B be sets, and $\pi_A \colon A \times B \to A$ the projection. (10)

For all α and β, the following diagram commutes:

$$\begin{array}{ccc} A \times B & \xrightarrow{\alpha \times \beta} & A \times B \\ \downarrow \pi_A & & \downarrow \pi_A \\ A & \xrightarrow{\alpha} & A \end{array} \tag{11}$$

Intuitively, by taking projections, we 'forget' about the effect of β. So, we can simply choose β such that all points in $A \times B$ are non-periodic with respect to β, while choosing α so that all points in A are periodic with respect to α. (12)

Let A be any set, $B = [0, 1]$, $\alpha = \mathrm{id}_A$ and $\beta : (x \mapsto \frac{1}{2}x)$ (13)

Wait, 0 is still a periodic point in the example above. Let's modify it. (14)

Let A be any nonempty set, $B = (0, 1]$, $\alpha = \mathrm{id}_A$ and $\beta : (x \mapsto \frac{1}{2}x)$ (15)

Clearly, β has no periodic points, so $\alpha \times \beta$ has no periodic points, but all points in A are periodic with respect to α. (16)

By (16), (2) follows. (17)

Exercise 1.2.3

Let G be a topological group. (1)

Prove that for each $g \in G$, the closure $H(g)$ of the set $\{g^n\}_{n=-\infty}^{\infty}$ is a commutative subgroup of G. (2)

Thus, if G has a minimal left translation, then G is abelian. (3)

Proof + reasoning:

First, let's show the closure of $\{g^n\}_{n=-\infty}^{\infty}$ is a subgroup of G, starting with showing closure under the group operation. (4)

Define $\langle g \rangle := \{g^n\}_{n=-\infty}^{\infty}$. (5)

Let $g \in G$. Let $a, b \in \mathrm{cl}(\langle g \rangle)$. (6)

What do I know about closures? The closure of A is the set X of points such that any neighborhood of $x \in X$ contains a point in A. (7)

My intuition is that the required proof will resemble the one I would follow if G were a metric space. In a metric space, if a and b are limits of g^{k_n} and g^{l_n} we should have $g^{k_n} g^{l_n} \to ab$. Here we do not have a metric, so there is no notion of convergent sequences, but instead of neighborhoods: x is a limit point of A if every neighborhood of x contains a point in A other than x itself. (8)

Intuitively, since a and b are in $\langle g \rangle$ or limit points of $\langle g \rangle$, the product of the two points that 'witness' this property should be the point that witnesses ab being a limit point. (9)

Let C be a neighborhood of ab, and $U \subseteq C$ an open set containing ab. (10)

$a^{-1}U$ and Ub^{-1} are open. (11)

$b \in a^{-1}U$ and $a \in Ub^{-1}$. Since a and b are limit points of $\langle g \rangle$, $\exists k, m \in \mathbb{Z}$ such that $g^m \in a^{-1}U$ and $g^k \in Ub^{-1}$. (12)

$g^k g^m$ should be in U, but I can't show why. What tools can I give myself to help prove $g^k g^m \in U$? (13)

Well, g^m and g^k are homeomorphisms, so $g^k g^m \in g^k a^{-1} U$, $g^k g^m \in U b^{-1} g^m$, and $(g^k a^{-1} U) \cup (U b^{-1} g^m)$ is open. (14)

Now I am stuck. (15)

What given assumptions have I not used? (16)

I have not used the fact that the group operation $G \times G \to G$ is continuous. I only used that, $\forall g \in G$, left and right multiplication by g is a continuous function $G \to G$, which seems to be a weaker statement. Using the 'joint' continuity should work. (17)

Since the group multiplication $\alpha : G \times G \to G$ is continuous, $\alpha^{-1}(U)$ is open in $G \times G$. Since $(a, b) \in \alpha^{-1}(U)$, and since sets of the form $A \times B$, where A and B are open, form a basis for the topology on $G \times G$, there exist open V and W such that $a \in V$, $b \in W$, and such that $V \times W \subseteq \alpha^{-1}(U)$. (18)

Since $a, b \in \mathrm{cl}(\langle g \rangle)$ there exist $g^\ell \in V$ and $g^p \in W$. By (12), $g^\ell g^p \in U$, hence $g^{\ell+p} \in U$, so $ab \in \mathrm{cl}(\langle g \rangle) = H(g)$. (19)

By (13), $H(g)$ is closed under taking products. (20)

Now we need to show that $H(g)$ has inverses, by showing $a^{-1} \in H(g)$. (21)

Let C be a neighborhood of a^{-1} and $U \subseteq C$ an open set such that $a^{-1} \in U$. Since the inverse is continuous, $U' := \{x \in G : x^{-1} \in U\}$ is open, and it contains a. (22)

Since $a \in H(g)$, there exists $g^q \in U'$, where $q \in \mathbb{Z}$. (23)

By (17), $g^{-q} = (g^q)^{-1} \in U$, so $a^{-1} \in H(g)$. (24)

By (19), $H(g)$ is closed under taking inverses. (25)

Now to prove that $H(g)$ is commutative. (26)

We need to show that $ab = ba$. If G were a metric space, the proof would follow from the fact that the limits of convergent sequences are unique. Is there something like uniqueness of limit points in a general topological space? The answer seems to be no, only when adding separation properties. (27)

Let's take a few steps back and try again. Note, the product in $H(g)$ is just the restriction of the one in G, so if $ab \neq ba$ in G, $ab \neq ba$ in $H(g)$. So, the only way in which $H(g)$ can be commutative is if it excludes at least all non-commutative elements in G. (28)

So, $H(g)$ must be a proper subgroup if G is not abelian. Considering (28), I think we should try to prove the contrapositive instead, i.e. prove if two elements of G are not commutative, then at least one of them is not in $H(g)$. (29)

Let $c, d \in G$ with $cd \neq dc$. (30)

Why is $(c, d) \notin H(g) \times H(g)$? I am stuck here. (31)

Why has my best attempt not worked? To show (31), we need to show that there exists a neighborhood of (c, d) containing no element of $\langle g \rangle$, but I can't find any obvious neighborhood. There is no given neighborhood from the definitions. I think the exercise is not correct without adding a separation property, so let's add it ourselves. (32)

Suppose that G is Hausdorff. (33)

By (16) and (30), there exist open neighborhoods U of cd and U' of dc such that $U \cap U' = \emptyset$. (34)

Suppose $c, d \in H(g)$. (35)

Similarly to (12), $(c, d) \in \alpha^{-1}(U)$ and $(d, c) \in \alpha^{-1}(U')$. (36)

So there are open sets V, V', W, W' such that $(c, d) \in V \times W \subseteq \alpha^{-1}(U)$ and $(d, c) \in V' \times W' \subseteq \alpha^{-1}(U')$. (37)

From (31), $c \in V \cap V'$ and $d \in W \cap W'$, and $V \cap V'$ and $W \cap W'$ are open. (38)

So, by (35), there exist $s, t \in \mathbb{Z}$ such that $g^s \in V \cap V'$ and $g^t \in W \cap W'$. (39)

By (24), $(g^s, g^t) \in V \times W$ and $(g^t, g^s) \in W' \times V'$. (40)

By (25) and (31), $g^s g^t \in U$ and $g^t g^s \in U'$, so $g^{t+s} \in U \cap U'$. (41)

(26) contradicts (16), so (35) is false, hence $c \notin H(g)$ or $d \notin H(g)$, so $H(g)$ is commutative. (42)

By (24), (14) and (20), $H(g)$ is a commutative subgroup of G. (43)

We still need to prove that if G has a minimal left translation, then G is Abelian. (44)

Suppose that G has a minimal left translation $L_h : G \to G$ where $h \in G$. (45)

By (28), $H(h)$ is a commutative subgroup of G. (46)

By definition, L_h has no proper closed non-empty invariant subsets. (47)

$H(h)$ is a closed non-empty subset of G. (48)

Is $H(h)$ invariant with respect to L_h? (49)

Let $a \in H(h)$. Let C be a neighborhood of ha and U open with $ha \in U \subseteq C$. $a \in h^{-1}U$, and $h^{-1}U$ is open, so $\exists q \in \mathbb{Z}$ such that $h^q \in h^{-1}U$. (50)

By (32), $h^{q+1} \in U$, so $H(h)$ is invariant. (51)

By (33), (30) and (33), $H(h) = G$, so G is abelian. (52)

Exercise 1.3.3

For $m \in \mathbb{Z}$, $|m| > 1$, define the times-m map $E_m : S^1 \to S^1$ by $E_m x = mx \mod 1$. Show that the set of points with dense orbits is uncountable. (1)

Proof + reasoning:

My first idea is to show that for any irrational x, the orbit under E_m is dense. How would we show this? We could use the semiconjugacy from (Σ_m, σ) to (S^1, E_m). As stated in ch1.3, the orbit of a point $0.x_1x_2\ldots$ is dense in S^1 iff every finite sequence of elements in $\{0, \ldots, m-1\}$ appears in the sequence $(x_i)_{i\in\mathbb{N}}$. (2)

Let's try proof by contradiction using (2). (3)

Let x be an irrational number. (4)

x has a base-m expansion $0.x_1x_2\ldots$ (5)

Suppose that the orbit of x is not dense. There exists a finite sequence $a_1 \ldots a_n$ of elements in $\{0, \ldots, m-1\}$ that does not occur anywhere in $x_1x_2\ldots$ (6)

Is the statement that any irrational x has a dense orbit true? Let's try a different approach. (7)

A more direct way to prove the statement is to construct an injective function from an uncountable set to the set of points in S^1 with dense orbits. It is noted in chapter 1.3 that we can construct a point in S^1 with dense orbit by simply concatenating all finite sequences. It seems likely that we can do something similar to construct the needed function. (8)

Let U be the set of points in S^1 with a unique base-m expansion. (9)

By the remarks in section 1.3, U is uncountable. (10)

Define $\phi : \Sigma_m \to S^1$. by $\phi((x_i)_{i\in\mathbb{N}}) := \sum_{i=1}^{\infty} x_i/m^i$ (11)

By the remarks in section 1.3, ϕ is bijective on $\phi^{-1}(U)$. (12)

Let $x \in U$, with base-m expansion $(x_i)_{i\in\mathbb{N}}$. (13)

Let $\mathcal{F}_m = \bigcup_{k=1}^{\infty}\{0, \ldots, m-1\}^k$. (14)

Clearly, $\mathcal{F}_m$ is countable, so it can be indexed by $(\omega_i)_{i\in\mathbb{N}}$. (15)

Define $\alpha : U \to \Sigma_m$ by letting $\alpha(x) = x_1\omega_1x_2\omega_2x_3\omega_3\ldots$, and define $\beta = \phi \circ \alpha$. (16)

Since every $y \in U$ has a unique base-m expansion, α is injective, so by (14), β is bijective. By construction, every finite sequence appears in $\alpha(y)$ for every $y \in U$, so by (2), every point in $\beta(U)$ has a dense orbit. (17)

From (19), (18), (10), we get that the set of all points in S^1 with dense orbits is uncountable. (18)

Exercise 1.4.3.

Verify that the metrics on Σ_m and Σ_m^+ generate the product topology (1)

Proof + reasoning:

Firstly, what product topology is being referred to? $\Sigma_m = A_m^{\mathbb{Z}}$, which is the space of sequences of elements in $\{1, \dots, m\}$ indexed by $\mathbb{Z}$, which can be seen as a product $\prod_{z \in \mathbb{Z}} A_m$. (2)

The product topology is generated by products of open sets. Is there a general result that countable product topologies are generated by cylinders, instead of just by countable products of open sets? (3)

Yes, the result is that the n-dimensional cylinders form a basis for the product topology. This is exactly the topology on Σ_m and Σ_m^+. (4)

How does (4) help us? From a metric we can generate the collection of open balls around all points. A metric generates a topology when, given a basis set B and any point $x \in B$, there is an open ball containing x that is contained in B. (5)

Let $C := C_{j_1,\dots,j_k}^{n_1,\dots,n_k} = \{x = (x_\ell) : x_{n_i} = j_i, i = 1, \dots, k\}$ where $n_1 < n_2 < \cdots < n_k$ are indices in $\mathbb{Z}$ or $\mathbb{N}$, and $j_i \in A_m$. (6)

Let $x := (x_i) \in C$. (7)

We want to show that C contains an open ball B such that $x \in B$. (8)

Recall that the metrics on Σ_m and Σ_m^+ are given by $d(x, x') = 2^{-l}$, where $l = \min\{|i| : x_i \neq x_i'\}$. If we pick ε small enough, we can construct $B(x, \varepsilon)$ such that all points in $B(x, \varepsilon)$ agree with x up to the 'largest' index in the definition of C. (9)

Let $m = \max\{|n_i| : i \leq k\}$,$y \in B(x, 2^{-m})$, and $l = \min\{|i| : y_i \neq x_i\}$. (10)

We have $2^{-l} = d(x, y) < 2^{-m}$. (11)

From (8), $l > m$, so $x_{n_i} = y_{n_i}$ $\forall i \leq k$, hence $y \in C$. (12)

Therefore $B(x, 2^{-m}) \subseteq C$. (13)

By (11), and the fact that the collection of sets such as $C_{j_1,\dots,j_k}^{n_1,\dots,n_k}$ form a basis for the product topology, the metrics generate the product topology. (14)

Exercise 1.5.3.

Suppose p is an attracting fixed point for f. Show that there is a neighborhood U of p such that the forward orbit of every point in U converges to p. (1)

Proof + reasoning:

Let's work out the meanings of the assumptions and of the required result. (2)

What is an attracting fixed point for f? A fixed point p is attracting if there exists a neighborhood U of p such that $\overline{U}$ is compact, $f(\overline{U}) \subseteq U$, and $\bigcap_{n\geq 0} f^n(U) = \{p\}$. (3)

From the definition of p, we are already given a candidate neighborhood U of p. Intuitively, it seems that $f^n(U)$ form a sequence of sets in which each is contained in the previous, with a single point p in the intersection, so orbits should converge to p or to a point in $\bigcap_{n\geq 0} f^n(\overline{U})$. But $f(\overline{U}) \subseteq U$ makes it likely that $\bigcap_{n\geq 0} f^n(\overline{U}) = \bigcap_{n\geq 0} f^n(U)$. Let's prove those intuitions. (4)

Is $f^{n+1}(U) \subseteq f^n(U)$ for all $n \in \mathbb{N}$? (5)

By assumption, there exists a neighborhood U of p such that $\overline{U}$ is compact, $f(\overline{U}) \subseteq U$, and $\bigcap_{n\geq 0} f^n(\overline{U}) = \{p\}$. (6)

Clearly,

$$U \subset \overline{U}. \tag{7}$$

From (6) and (5)

$$f(U) \subseteq f(\overline{U}) \subseteq U. \tag{8}$$

Therefore,

$$f^{n+1}(U) \subseteq f^n(U) \text{ for all } n \in \mathbb{N} \tag{9}$$

Is it true that $\bigcap_{n\geq 0} f^n(\overline{U}) = \bigcap_{n\geq 0} f^n(U)$? (10)

Clearly,

$$\bigcap_{n\geq 0} f^n(U) \subseteq \bigcap_{n\geq 0} f^n(\overline{U}) \tag{11}$$

Conversely,

$$\bigcap_{n\geq 0} f^n(\overline{U}) \subseteq \bigcap_{n\geq 1} f^n(\overline{U}) = \bigcap_{n\geq 0} f^{n+1}(\overline{U}) \subseteq \bigcap_{n\geq 0} f^n(U) \tag{12}$$

So, from (12)

$$\bigcap_{n\geq 0} f^n(U) = \bigcap_{n\geq 0} f^n(\overline{U}) \tag{13}$$

The answer to (10) is yes. Now, let's try to prove convergence of forward orbits. (14)

Let $x \in U$. Define $(x_n)_{n\in\mathbb{N}} = (f^n(x))_{n\in\mathbb{N}}$. (15)

Does (x_n) converge to p? (16)

Let's try proof by contradiction. (17)

Assume (x_n) does not converge. Then $\exists \varepsilon' > 0$ such that $\forall n : \exists k \geq n : d(f^k(x), p) > \varepsilon'$. (18)

I'm stuck. I know there are infinitely many points in $U \setminus B(p, \varepsilon')$. Intuitively, this should contradict the fact that $\bigcap_{n\geq 0} f^n(\overline{U})$ only contains p. (19)

Let me check what given assumptions I haven't used in my proof yet. I haven't used that p is a fixed point, that f is continuous, or the compactness of $\overline{U}$. (20)

Let me think of properties that we can use, or look up general consequences of these facts. (21)

In a metric space, the intersection of compact sets is compact. (22)

If some $y \in U$ is not in $f^n(U)$ for some $n \in \mathbb{N}$, it won't be in any $f^k(U)$ for $k \geq n$. (23)

In a compact metric space, all sequences have convergent subsequences. (24)

I think I can apply (20) to show a contradiction: (25)

From (17), there exists a sequence $f^{m_n}(x)$ such that $d(f^{m_n}(x), p) \geq \varepsilon$ for all $n \geq 0$. By compactness of $\overline{U}$, this sequence has a convergent subsequence $(f^{z_n}(x))_{n \geq 0}$ with $f^{z_n}(x) \to z \in \overline{U}$ and $z_n \to \infty$. (26)

Since f is continuous and $\overline{U}$ compact, $f^n(\overline{U})$ is compact for all $n \geq 0$. (27)

$\forall n \geq 0$ there exists K s.t. $f^{z_K}(x) \in f^n(\overline{U})$, hence $\forall m \geq K$, $f^{z_m}(x) \in f^n(\overline{U})$. (28)

From (24), the limit point z must be in $f^n(\overline{U})$ for all $n \geq 0$. (29)

Therefore, $z \in \bigcap_{n \geq 0} f^n(\overline{U}) = \{p\}$. (30)

So, $z = p$, which contradicts (17), so assumption (17) is false. Hence (x_n) converges. (31)

We still need to show that (x_n) converges to p, but the proof will likely be almost the same as above. (32)

Suppose $x_n \to q$ and $q \neq p$. (33)

Let $n \geq 0$. The sequence $(x_i)_{i \geq n}$ is contained in $f^n(\overline{U})$, which is compact, so $q \in f^n(\overline{U})$. Therefore, $q \in \bigcap_{n \geq 0} f^n(\overline{U})$. (34)

By (13), $q \in \bigcap_{n \geq 0} f^n(U) = \{p\}$, which is a contradiction. Hence, (17) is false. (35)

From (31) and (16) it follows that $x_n \to p$. Therefore, the forward orbit of any point in U converges to p. (36)

Exercise 1.7.3.

Show that the eigenvalues of a two-dimensional hyperbolic toral automorphism are irrational (so the stable and unstable manifolds are dense by exercise 1.11.1). (1)

Proof + reasoning:

Are there hyperbolic toral automorphisms that aren't represented by a matrix? No, not in this context. (2)

Let A be a 2×2 integer matrix such that $\det(A) = 1$ and such that for all eigenvalues λ of A, $|\lambda| \neq 1$. (3)

Let λ be an eigenvalue of A. (4)

Let's try the following: first, relate the eigenvalues to the determinant. Then, conclude from the first step and the given assumptions that the eigenvalues are irrational. (5)

Note that $\det(\lambda I - A) = 0$. (6)

Denote $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. (7)

$(\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21} = 0$ (8)

$\lambda^2 - \lambda a_{11} - \lambda a_{22} + a_{11}a_{22} - a_{12}a_{21} = 0$ (9)

By assumption, $\det(A) = 1$, so $a_{11}a_{22} - a_{12}a_{21} = 1$ (10)

Substituting $1 = a_{11}a_{22} - a_{12}a_{21}$ in (10) gives: $\lambda^2 - \lambda a_{11} - \lambda a_{22} + 1 = 0$ (11)

$\lambda^2 - \lambda(a_{11} + a_{22}) + 1 = 0$ (12)

That was the first step. How can we conclude? (13)

What given assumption have I not used yet? (14)

I haven't used that $|\lambda| \neq 1$. From (13) it may be possible to conclude that λ is equal to 1 or irrational, which together with (15) gives that λ is irrational. (15)

We can factorize the left-hand side of (13). (16)

However, (13) could have complex solutions for certain values of $a_{11} + a_{22}$. (17)

If the discriminant of (13) is greater than 0, (13) only has real roots. (18)

I'm not sure how to proceed. Let's take a few steps back. (19)

Equation (13) is quadratic in λ, so all solutions are given by the quadratic formula. (20)

By the quadratic formula, $\lambda = \frac{1}{2}(a_{11} + a_{22} + ((a_{11} + a_{22})^2 - 4)^{1/2})$ or $\lambda = \frac{1}{2}(a_{11} + a_{22} - ((a_{11} + a_{22})^2 - 4)^{1/2})$ (21)

From (14) we see that $\lambda \in \mathbb{R} \setminus \mathbb{Q}$ if and only if $((a_{11} + a_{22})^2 - 4)^{1/2} \in \mathbb{R} \setminus \mathbb{Q}$ (22)

By definition of A, $(a_{11} + a_{22}) \in \mathbb{N}$ (23)

Let's show that λ cannot be in $\mathbb{C}$. My guess is that if λ were in $\mathbb{C}$, its magnitude would be equal to 1. (24)

Suppose $((a_{11} + a_{22})^2 - 4)^{1/2} \in \mathbb{C}$. Then $|\lambda|^2 = \frac{1}{4}((a_{11} + a_{22})^2 + 4 - (a_{11} + a_{22})^2) = 1$ (25)

(25) contradicts (4), hence $((a_{11} + a_{22})^2 - 4)^{1/2} \in \mathbb{R}$ (26)

From (19) and (16), $a_{11} + a_{22} \geq 3$ (27)

We just need to show that $\lambda \notin \mathbb{Q}$. My intuition is that $((a_{11} + a_{22})^2 - 4)^{1/2}$ is always irrational, because it seems $\forall n \in \mathbb{N}$, $n^{1/2}$ is rational only if n is a square number, and subtracting 4 makes it no longer square, given $\sqrt{n} \geq 3$. Let's prove the first part. (28)

Conjecture: $\{n \in \mathbb{N} : \sqrt{n} \in \mathbb{Q} \setminus \mathbb{N}\} = \emptyset$ (29)

Proof of (29): Suppose $\sqrt{n} \in \mathbb{Q} \setminus \mathbb{N}$. Then $\sqrt{n} = \frac{p}{q}$, where p and q are natural numbers. (30)

Since $\sqrt{n} \notin \mathbb{N}$, q does not divide p, so q^2 does not divide p^2 (31)

From (30), $n = \frac{p^2}{q^2}$, hence $nq^2 = p^2$ (32)

(27) contradicts (17), so (30) is false $\square_{c23}$ (33)

Now let's prove the second part, that subtracting 4 makes the number no longer square. Intuitively, this seems true because the distance between consecutive squares will eventually be greater than any fixed number, like 4, so by subtracting a fixed number from large enough squares, we end up in between squares. (34)

Conjecture: $\forall n \in \mathbb{N}, n \geq 3$ implies $(n^2 - 4)^{1/2} \notin \mathbb{N}$ (35)

Proof of (35): Suppose $(n^2 - 4)^{1/2} = k$ where $k \in \mathbb{N}$ (36)

Then $n^2 - 4 = k^2$ (37)

So $n^2 - k^2 = 4$ (38)

Clearly, $n > k$ (39)

Then $n^2 - k^2 \geq n^2 - (n-1)^2 = n^2 - n^2 + 2n - 1 = 2n - 1$ (40)

So $n^2 - k^2 > 4$ (41)

(29) contradicts (25), so $\square_{c34}$ (42)

Now we can conclude. By (20) and (35), $((a_{11} + a_{22})^2 - 4)^{1/2} \notin \mathbb{N}$ (43)

So, by (30), (19) and (29), $((a_{11} + a_{22})^2 - 4)^{1/2}$ is irrational (44)

So, by (31) and (14), λ is irrational (45)

We still need to show that the (un)stable manifolds are dense by exercise 1.11.1. (46)

Let $x \in \mathbb{T}^2$. Without loss of generality, $\lambda > 1 > \lambda^{-1}$, where λ and λ^{-1} are the eigenvalues of A. The stable manifold $W^u(x)$ is the line through x parallel to v where v is the eigenvector corresponding to λ. (47)

I think the intended proof is to show that the slope of a line parallel to v is irrational, hence the flow defined in section 1.11 has dense orbits, from which it follows $W^u(x)$ is dense. (48)

Denote $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. A line parallel to v has slope equal to $\frac{v_2}{v_1}$ (49)

My guess is that the slope is irrational because λ is irrational and because v is an eigenvector of an integer-valued matrix. Let's write out the defining equation for eigenvectors. (50)

$Av = \lambda v$ (51)

$$\begin{bmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{bmatrix} = \begin{bmatrix} \lambda v_1 \\ \lambda v_2 \end{bmatrix} \quad (52)$$

$a_{11}v_1 + a_{12}v_2 = \lambda v_1$ (53)

$a_{21}v_1 + a_{22}v_2 = \lambda v_2$ (54)

$v_2(\lambda - a_{22}) = a_{21}v_1$ (55)

$v_1(\lambda - a_{11}) = a_{12}v_2$ (56)

If $a_{21} = 0$, then $v_2 v_1^{-1} = 0$. Why is this not possible? Probably because v is nonzero, by assumption. (57)

Since v is an eigenvector, $v_1 \neq 0$ or $v_2 \neq 0$ (58)

By (32), $(\lambda - a_{22})$ and $(\lambda - a_{11})$ are irrational. (59)

If $v_1 \neq 0$, then by (59) and (28), $v_2 \neq 0$ (60)

If $v_2 \neq 0$, then by (59) and (38), $v_1 \neq 0$ (61)

So, $v_1 \neq 0$ and $v_2 \neq 0$ (62)

By (62) and (38), $v_2 v_1^{-1} = (\lambda - a_{22})^{-1} a_{21} \neq 0$ (63)

By (63) and (59), $v_2 v_1^{-1}$ is irrational (64)

Denote $x = (x_1, x_2)$, let $t \in \mathbb{R}^+$ (65)

Let $\phi^t_{\frac{v_1}{v_2}}(x) := (x_1 + \frac{v_1}{v_2}t, x_2 + t) \mod 1$ (66)

Then $\bigcup_{t \in \mathbb{R}^+} \phi^t_{\frac{v_1}{v_2}}(x) \subseteq W^u(x)$ (67)

By exercise 1.11.1, the orbit of $\phi_{\frac{v_1}{v_2}}$ is dense, so $W^u(x)$ is dense. For the stable manifold, the proof is similar. (68)

Exercise 1.8.3.

Let $\phi : \Sigma_2 = \{0,1\}^{\mathbb{Z}} \to H$ be the map that assigns to each infinite sequence $\omega = (\omega_i) \in \Sigma_2$ the unique point $\phi(\omega) = \bigcap_{-\infty}^{\infty} f^{-i}(R_{\omega_i})$. (1)

Prove that ϕ is a bijection and that both ϕ and ϕ^{-1} are continuous. (2)

Proof + reasoning:

We will show ϕ is injective. (3)

Suppose $x, y \in \Sigma_2$ with $\phi(x) = \phi(y)$. Then

$$\bigcap_{-\infty}^{\infty} f^{-i}(R_{x_i}) = \phi(x) = \phi(y) = \bigcap_{-\infty}^{\infty} f^{-i}(R_{y_i}). \tag{4}$$

Is f injective? It seems so, since f is a map that stretches and bends the space D into a horseshoe, none of the regions of the horseshoe seem to overlap, and ϕ is a conjugacy. (5)

From the description of f, we see that f is injective. (6)

By definition,

$$R_0 = f(D_0) \cap R \quad \text{and} \quad R_1 = f(D_1) \cap R. \tag{7}$$

From (6), (4), and $D_0 \cap D_1 = \emptyset$, we get $R_0 \cap R_1 = \emptyset$. (8)

By (6) and (5), $f^{-i}(R_0) \cap f^{-i}(R_1) = \emptyset$ for all $i \in \mathbb{Z}$. (9)

From (2) and (6), $x_i = y_i$ for all $i \in \mathbb{Z}$, so $x = y$, so ϕ is injective. (10)

Let's show ϕ is surjective. (11)

Note $f(R) \cap R = R_1 \cup R_0$, and $f^{-1}(R) \subseteq R$. (12)

By (15), $f^{-i}(R) = f^{-i}(R_0) \cup f^{-i}(R_1)$ for all $i \geq 1$. (13)

Clearly, $R_0 \cap R_1 = \emptyset$. (14)

By (6) and (15), for all $i \geq 0$,

$$\left(f^{-i}(R_0) \cup f^{-i}(R_1)\right) \cap R = f^{-i}(R_0 \cup R_1) \cap R = f^{-i+1}(R) \cap f^{-i}(R) \cap R. \tag{15}$$

Let $x \in H$ and $j \in \mathbb{Z}$. By (16), (17) and (18), $x \in f^j(R_0)$ or $x \in f^j(R_1)$, but not both. (16)

By (12), we can define $x_j = 0$ if $x \in f^j(R_0)$ and $x_j = 1$ if $x \in f^j(R_1)$. Clearly, this gives a sequence $(x_j)_{j\in\mathbb{Z}} \in \Sigma_2$ such that $\phi((x_j)_{j\in\mathbb{Z}}) = x$. So ϕ is surjective. (17)

By (17) and (7), ϕ is bijective. (18)

Next, we show that ϕ is continuous. It seems that sets of the form $f^{-i}(R_{\omega_i}) \times R_{\omega_{i+1}}, \ldots, R_{\omega_n}$ are open, and even form a basis for the topology. Proving this will simplify the remaining parts of the exercise, since it suffices to prove that the inverse images of basis sets are open. (19)

Let $\omega \in \phi^{-1}(A \times B)$. For a sequence $\omega \in \{0,1\}^{\mathbb{Z}}$, define

$$R_{\omega_{-m},\ldots,\omega_m} = \bigcap_{i=-m}^{m} f^{-i}(R_{\omega_i}). \tag{20}$$

Define $\mathcal{C}_m = \{R_{\omega_{-m},\ldots,\omega_m} \times R_{\omega_0,\ldots,\omega_m}\}$, $\omega \in \{0,1\}^{\mathbb{Z}}$, $m \in \mathbb{N}$, and define $\mathcal{C} = \bigcup_{m\in\mathbb{N}}\{H \cap C : C \in \mathcal{C}_m\}$. (21)

Conjecture 1. *$\mathcal{C}$ is a basis for the topology on H.*

Proof: Let $C \in \mathcal{C}$. Then $C = H \cap (R^- \times R^+)$ where $R^- \times R^+ \in \mathcal{C}_m$ for some $m \in \mathbb{N}$. (22)

Is C open? Intuitively, yes, since the sets in $\mathcal{C}_m$ are closed and bounded, we can contain them in open sets in $\mathbb{R}^2$ each containing no other points from H. (23)

Note $R^- = [x_1, x_2]$ and $R^+ = [y_1, y_2]$ for $x_1, x_2, y_1, y_2 \in \mathbb{R}$. (24)

By (17), for all $D \neq D' \in \mathcal{C}_m$, $D \cap D' = \emptyset$. (25)

By (24) and (28), there exist open intervals $I^-, I^+ \subset \mathbb{R}$ such that $R^- \times R^+ \subseteq I^- \times I^+$ and such that $(I^- \times I^+) \cap D = \emptyset$ for all $D \in \mathcal{C}_m$ with $D \neq R^- \times R^+$. (26)

Clearly, $\mathcal{C}_m$ covers H. (27)

By (29) and (27), $H \cap (R^- \times R^+) = H \cap (I^- \times I^+)$. (28)

By (28), C is open in H. (29)

Aren't we done at this point, since H is covered by $\mathcal{C}_m$ for all m? No, we still need to check that we can build each open U out of sets in $\mathcal{C}$ that are contained in U. Intuitively, this seems true because the rectangles in $\mathcal{C}_m$ get arbitrarily small as m increases. Let's prove this. (30)

Let A and B be open intervals in $\mathbb{R}$. (31)

Let $x \in H \cap (A \times B)$. (32)

Let $\varepsilon = \min\{d(x, y) : y \in A \times B\}$. (33)

Because $A \times B$ is open, $\varepsilon > 0$. (34)

Let $k = \min\{n \in \mathbb{N} : \mu^{-n} \leq \varepsilon, \lambda^n \leq \varepsilon\}$. (35)

By (30), and since $\lambda < 1/2$ and $\mu > 2$, $k > 0$. (36)

If $R^- \times R^+ \in \mathcal{C}_m$, then R^- has width equal to $\mu^{-k} \leq \varepsilon$ and R^+ has width equal to $\lambda^k \leq \varepsilon$. (37)

By (37), (33) and (27), there exists an $R^- \times R^+ \in \mathcal{C}_k$ such that $x \in R^- \times R^+$ and $R^- \times R^+ \subseteq A \times B$. (38)

By (29), $H \cap R^- \times R^+$ is open, so $\mathcal{C}$ is a basis for the topology on H. (39)

We can use 2 to prove ϕ is continuous, since we just need to prove $\phi^{-1}(C)$ is open for $C \in \mathcal{C}$. (40)

Let $C \in \mathcal{C}$. $C = H \cap (R_{\omega_{-m},\ldots,\omega_m} \times R_{\omega_0,\ldots,\omega_m})$ for $\omega \in \{0,1\}^{\mathbb{Z}}$. (41)

Let's show $\phi^{-1}(C) = B(\omega, 2^{-m})$. (42)

Let $j \in \{-m, -m+1, \ldots, m\}$. (43)

Suppose $z \in \phi^{-1}(C)$. By definition of ϕ,

$$\phi(z) = \bigcap_{i\in\mathbb{Z}} f^{-i}(R_{z_i}) \subseteq f^j(R_{z_j}). \tag{44}$$

Since $\phi(z) \in R^+ \times R^-$, $\phi(z) \in f^j(R_{\omega_j})$. (45)

By (17), $f^j(R_1) \cap f^j(R_0) = \emptyset$, so $z_j = \omega_j$, so $z \in B(\omega, 2^{-m})$. (46)

Clearly $B(\omega, 2^{-m}) \subseteq \phi^{-1}(C)$. (47)

By (41) and (42), $\phi^{-1}(C) = B(\omega, 2^{-m})$, so $\phi^{-1}(C)$ is open, so, by 2, ϕ is continuous. (48)

Now let's show ϕ^{-1} is continuous. (49)

Let $B(\gamma, 2^{-n})$ be an open ball in Σ_2. By the same argument as for (48),

$$B(\gamma, 2^{-n}) = \phi^{-1}(H \cap (R_{\gamma_{-n},\dots,\gamma_n} \times R_{\gamma_0,\dots,\gamma_n})). \tag{50}$$

So $\phi(B(\gamma, 2^{-n})) = H \cap (R_{\gamma_{-n},\dots,\gamma_n} \times R_{\gamma_0,\dots,\gamma_n})$. (51)

By (46) and 2, $\phi(B(\gamma, 2^{-n}))$ is open, so ϕ^{-1} is continuous. (52)

Exercise 1.9.3.

Let $\mathbb{T}$ denote the set of sequences $(\phi_i)_{i=0}^{\infty}$ where $\phi_i \in S^1$ and $\phi_i = 2\phi_{i+1} \mod 1$ for all i. Let $\alpha : \mathbb{T} \to \mathbb{T}$ be defined by

$$(\phi_0, \phi_1, \dots) \mapsto (2\phi_1, \phi_1, \phi_2, \dots). \tag{1}$$

Show that $\mathbb{T}$ is a topological group. (2)

Show that α is an automorphism (3)

Proof + reasoning:

Let's first formulate basic lemmas to use throughout the exercise. (4)

Lemma 1. *$\forall x \in \mathbb{R}, \forall y \in \mathbb{R},\ a, b \in \mathbb{N} : x \equiv_a y \Rightarrow bx \equiv_a by$.*

Proof. $\exists p \in \mathbb{Z} : x = y + pa$, so $bx = by + bpa$, so $bx \equiv_a by$. □

Lemma 2. *$\forall x \in \mathbb{R}, a, b \in \mathbb{N},\ a(x \mod b) \equiv_b ax$.*

Proof. Clearly, $x \mod b \equiv_b x$. By lemma 1, $a(x \mod b) \equiv_b ax$. □

Lemma 3. *$\forall x, y \in \mathbb{R},\ \forall k \in \mathbb{N},\ (x \mod k) + (y \mod k) \equiv_k x + y$.*

Proof. $x + y = (x \mod k) + (y \mod k) + pk + qk = (x \mod k) + (y \mod k) + (p+q)k$. So, $x + y \equiv_k (x \mod k) + (y \mod k)$. □

Given ψ and ϕ in $\mathbb{T}$, define $(\psi + \phi)_i = (\psi_i + \phi_i) \mod 1$. (5)

Let ψ and ϕ be elements of $\mathbb{T}$. It suffices to show $\psi_i + \phi_i \equiv_1 2(\psi + \phi)_{i+1}$. (6)

By (1), lemma 3 and lemma 2,

$$\begin{aligned}
\psi_i + \phi_i &= (2\psi_{i+1} \mod 1) + (2\phi_{i+1} \mod 1) \\
&\equiv_1 2\psi_{i+1} + 2\phi_{i+1} \\
&= 2(\psi_{i+1} + \phi_{i+1}) \\
&\equiv_1 2((\psi + \phi)_{i+1} \mod 1) \\
&\equiv_1 2(\psi + \phi)_{i+1}
\end{aligned} \tag{7}$$

From (7) and (6), $\mathbb{T}$ is closed under addition. (8)

For $\phi \in \mathbb{T}$, define $(-\phi)_i := -\phi_i$. (9)

Clearly, this is the inverse of ϕ. (10)

Now, let's check continuity. (11)

Suppose that $\phi^n \to \phi$ and $\psi^n \to \psi$ in $(S^1)^{\mathbb{N}_0}$. Let $n \geq 0$. By assumption, $\exists K_n$ s.t. $\forall j \leq n,\ \phi^i_j = \phi_j,\ \psi^i_j = \psi_j$, hence $(\phi^i - \psi^i)_j = \phi_j - \psi_j$. (12)

By (12), $\forall \varepsilon > 0, \exists k$ s.t. $d(\phi^i - \psi^i, \phi - \psi) \leq \varepsilon\ \forall i \geq k$, so $\lim_{n\to\infty} d(\phi^n - \psi^n, \phi - \psi) = 0$, hence $(\phi, \psi) \mapsto \phi - \psi$ is continuous. (13)

Now, we check α is a group automorphism. Is it possible to take a shortcut? I think that if α preserves the product and identity and is bijective, then it follows that α preserves inverses and that α^{-1} is a group homomorphism. (14)

Let $\phi, \psi \in \mathbb{T}$. Show $\alpha(\phi) + \alpha(\psi) = \alpha(\phi + \psi)$. By lemma 3 and lemma 2,

$$\begin{aligned}(\alpha(\phi) + \alpha(\psi))_i &= ((2\phi_i) \bmod 1 + (2\psi_i) \bmod 1) \bmod 1 \\ &\equiv_1 2(\phi_i + \psi_i) \\ &\equiv_1 2\left((\phi + \psi)_i \bmod 1\right) \\ &= \alpha(\phi + \psi)_i.\end{aligned} \tag{15}$$

Clearly, α preserves the identity. (16)

Now, let's show α is bijective. (17)

Suppose $\alpha(\phi) = \alpha(\psi)$. Then $\phi_{i-1} = \alpha(\phi)_i = \alpha(\psi)_i = \psi_{i-1}$ $\forall i \geq 1$, so $\phi = \psi$, hence α is injective. (18)

Let $\phi \in \mathbb{T}$. Let $(\phi')_i := \phi_{i+1}$ $\forall i \geq 1$. (19)

$\alpha(\phi') = \phi$, so α is surjective. (20)

By (18) and (20), α is bijective. By (15), (16), and (18), α is a group automorphism. (21)

Now we need to show α is a homeomorphism. Note that α is a map between product topologies. This should simplify our proofs. (22)

For product topologies the 1-d cylinders form a subbasis. So, to show that α is continuous it suffices to show that $\forall i \in \mathbb{N}, \pi_i \circ \alpha$ is continuous. (23)

Let $i \in \mathbb{N}$. Note $\pi_i \circ \alpha : (\phi_0, \phi_1, \dots) \mapsto \begin{cases} \phi_{i-1} & \text{if } i \geq 1 \\ 2\phi_0 & \text{if } i = 0 \end{cases}$. (24)

The map $r : S^1 \to S^1 : s \mapsto 2s \mod 1$ is clearly continuous. (25)

By (25), if $A \in \mathcal{T}(S^1)$, then

$$(\pi_i \circ \alpha)^{-1}(A) = \begin{cases} (S^1)^{i-1} \times A \times S^1 \times \cdots & \text{if } i \geq 1 \\ r^{-1}(A) \times S^1 \times \cdots & \text{if } i = 0 \end{cases} \tag{26}$$

By (26) and (25), $\pi_i \circ \alpha$ is continuous, so by (23), α is continuous. (27)

Note $\pi_i \circ \alpha^{-1} : (\phi_0, \phi_1, \dots) \mapsto \phi_{i+1}$. This is clearly continuous, so by (23), α^{-1} is continuous, so by (27) and (21), α is a homeomorphism. (28)

Exercise 1.10.3.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. (1)

Show that $-f$ is a Lyapunov function for the gradient flow. (2)

Show that the trajectories of the gradient flow are orthogonal to the level sets of f. (3)

Proof + reasoning:

Let's write out the relevant definitions. (4)

The gradient flow is the flow of the differential equation $\dot{x} = \nabla f(x)$. (5)

Denote the time-t gradient flow by $g^t : \mathbb{R}^n \to \mathbb{R}^n$. For all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}^+$, write $g_x(t) := g^t(x)$. (6)

For each x this defines $g_x : \mathbb{R}^+ \to \mathbb{R}^n$. (7)

Let $x \in \mathbb{R}^n$ and $t \in \mathbb{R}^+$. Note $(f \circ g_x)(0) = f(x)$ and $(f \circ g_x)(t) = f(g^t(x))$. (8)

By (7), if $(f \circ g_x)'(s) \geq 0$ for all $s \in \mathbb{R}^+$ then $-f$ is Lyapunov. (9)

By (4), $g_x'(t) = \nabla f(g_x(t))$. (10)

By the multivariate chain rule and (10),

$$(f \circ g_x)'(t) = \langle \nabla f(g_x(t)), g_x'(t) \rangle = \langle g_x'(t), g_x'(t) \rangle$$

where $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^n$. (11)

By definition of inner products, $\langle g_x'(t), g_x'(t) \rangle \geq 0$. (12)

By (11), (13) and (8), $-f$ is Lyapunov. (13)

Next, we want to show statement (2). (14)

To express orthogonality, we need a common inner product space, but this is just $\mathbb{R}^n$ in our case. (15)

Which vectors are we trying to prove are orthogonal? I think, given some point $x \in \mathbb{R}^n$, we should compare the time derivative of the orbit of x at $t = 0$, with a vector in $\mathbb{R}^n$ 'tangent' to the level set of f at $f(x)$. (16)

How can we define the tangent vector? (17)

Let $x \in \mathbb{R}^n$. (18)

Define the level set $C := f^{-1}(f(x))$. (19)

C is a subset of $\mathbb{R}^n$, but I don't think it is necessarily a smooth manifold. Still, we can define tangent vectors in terms of smooth paths in $\mathbb{R}^n$: (20)

Let $T_x = \{\dot{\gamma}(0) : \exists \varepsilon > 0 \text{ s.t. } \gamma : (-\varepsilon, \varepsilon) \to C \text{ is smooth and } \gamma(0) = x\}$. (21)

By (10), $(f \circ g_x)'(0) = \nabla f(g_x(0)) = \nabla f(x) = g_x'(0)$. (22)

Let $V \in T_x$, with corresponding path $\gamma : (-\varepsilon, \varepsilon) \to C$. (23)

We need to show $\langle V, \nabla f(x) \rangle = 0$. (24)

Since $\gamma(t) \in C$ for all $t \in (-\varepsilon, \varepsilon)$, $f(\gamma(t)) = f(\gamma(0))$ for all $t \in (-\varepsilon, \varepsilon)$. (25)

By the multivariate chain rule,

$$\begin{aligned}
\langle \nabla f(x), V\rangle &= \sum_{k=1}^{n} V_k \frac{\partial f}{\partial y_k}(x) \\
&= \sum_{k=1}^{n} V_k \frac{\partial f}{\partial y_k}(\gamma(0)) \\
&= \sum_{k=1}^{n} \dot{\gamma}(0)_k \frac{\partial f}{\partial y_k}(\gamma(0)) \\
&= (f \circ \gamma)'(0). \qquad (26)
\end{aligned}$$

By (25), $(f \circ \gamma)'(0) = 0$, so by (16), $\langle \nabla f(x), V\rangle = 0$, so by (22), the trajectories of the gradient flow are orthogonal to the level sets of f. (27)

Exercise 1.11.3.

Suppose $1, s$ and αs are real numbers that are linearly independent over $\mathbb{Q}$. (1)

Show that every orbit of the time-s map ϕ_α^s is dense in $\mathbb{T}^2$. (2)

Proof + reasoning:

Let's try to adapt the proof that R_α has dense semiorbits if α is irrational. (3)

The idea in (3) is to divide S^1 into disjoint ε-sized intervals, and use the pigeonhole principle to show that there exist $k > m$ such that $R_\alpha^k(x)$ and $R_\alpha^m(x)$ belong to the same interval, hence that R_α^{k-m} is a translation by less than ε. In our case, we need to use rectangles instead of intervals, and the translation is not simply defined as the addition of an irrational number. (4)

Let $x \in \mathbb{T}^2$, $y \in \mathbb{T}^2$, $\varepsilon' > 0$ and $\varepsilon = \frac{\varepsilon'}{2\sqrt{2}}$. (5)

Let $\mathcal{P}_\varepsilon$ be a partition of $\mathbb{T}^2$ into finitely many squares of the form $[a,b)^2$, where $\frac{\varepsilon}{2} < |a-b| < \varepsilon$. (6)

By the pigeonhole principle, there exists a $P \in \mathcal{P}_\varepsilon$ and $k > m$ in $\mathbb{Z}$ such that $\phi_\alpha^{ks}(x)$ and $\phi_\alpha^{ms}(x)$ are in P. (7)

By (6), $d(z, \phi_\alpha^{(k-m)s}(z)) < \sqrt{2}\varepsilon$ for all $z \in \mathbb{T}^2$, where d is the metric on $\mathbb{T}^2$. (8)

Intuitively, $\phi_\alpha^{(k-m)s}$ is a translation of size less than $\sqrt{2}\varepsilon$, seemingly with an irrational slope. So, if the line from x in the direction of that translation eventually intersects an ε-ball around y, then there should exist an iterate of x that is within ε' of y. Let's make this precise. We can state the conjecture first, and prove it only after we know that it gives the required result. (9)

Conjecture 2. *There exists a $\beta \in \mathbb{R} \setminus \mathbb{Q}$ such that for all $y \in \mathbb{T}^2$*

$$\frac{(\phi_\alpha^{(k-m)s}(y))_2 - y_2}{(\phi_\alpha^{(k-m)s}(y))_1 - y_1} = \beta.$$

Proof. Suppose for contradiction that $s = 0$. Then for $p = 1, q = 1, r = 0$ we have $p\alpha s + qs + r = 0$, a contradiction, so $s \neq 0$. Similarly, $\alpha s \neq 0$. Suppose for contradiction that $\alpha s \in \mathbb{Q}$. Let $p = 1, q = -\alpha s, r = 0$. Then $p\alpha s + qs + r = 0$, a contradiction, so $\alpha s \notin \mathbb{Q}$. Suppose for contradiction that $\frac{1}{\alpha} \in \mathbb{Q}$. Then s is irrational. Let $p = \frac{1}{\alpha}, q = -1, r = 0$. Then $p\alpha s + qs + r = 0$, a contradiction, so $\frac{1}{\alpha}$ is irrational. Let $y \in \mathbb{T}^2$. Then

$$\frac{(\phi_\alpha^{(k+m)s}(y))_2 - y_2}{(\phi_\alpha^{(k+m)s}(y))_1 - y_1} = \frac{(k-m)s}{(k-m)\alpha s} = \frac{1}{\alpha}$$

So, with $\beta = \frac{1}{\alpha}$, the statement follows. □

Let γ be the line in $\mathbb{T}^2$ starting from x in the direction of $x - \phi_\alpha^{m-k}(x)$. (10)

Let β be the slope of γ, which is finite and in $\mathbb{R} \setminus \mathbb{Q}$ by 2. (11)

By (13), considering γ as a subset of $\mathbb{T}^2$, we have

$$\begin{aligned}\gamma \cap (y_1 \times \mathbb{T}) &= \bigcup_{n \geq 0} \{(y_1, (x_2 + \beta(y_1 - x_1) + \beta n) \bmod 1)\} \\ &= \bigcup_{n \geq 0} \left\{ \left(y_1, R_\beta^n(x_2 + \beta(y_1 - x_1))\right) \right\}.\end{aligned} \tag{12}$$

By (13), R_β has dense semiorbits. (13)

By (13) and (12), there exists a $z \in \gamma \cap (y_1 \times (y_2 - \varepsilon, y_2 + \varepsilon))$. (14)

By (8) and Conjecture 2, there exists a $p \in \mathbb{N}$ such that

$$d(\phi_\alpha^{p(k-m)s}(x), z) < \sqrt{2}\varepsilon \tag{15}$$

By (15) and (14),

$$\begin{aligned}d(\phi_\alpha^{p(k-m)s}(x), y) &\leq d(\phi_\alpha^{p(k-m)s}(x), z) + d(z, y) \\ &\leq \sqrt{2}\varepsilon + \varepsilon \\ &\leq 2\sqrt{2}\varepsilon \\ &\leq \varepsilon'.\end{aligned} \tag{16}$$

By (16), every orbit of ϕ_α^s is dense in $\mathbb{T}^2$. (17)

Exercise 1.12.3.

Compute the Lyapunov exponents for the solenoid. (1)

Proof + reasoning:

We need to calculate the matrix corresponding to the total derivative of F^n. (2)

Note the total derivative of $f\colon \mathbb{R}^n \to \mathbb{R}^n$ corresponds to the matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{3}$$

We need to compute $dF^n(x)v$. Is there a shortcut that we can take? Since we are calculating the derivative of a composition of functions, we might use the chain rule. (4)

The chain rule for total derivatives states $d(f \circ g)(a) = df(g(a)) \circ dg(a)$. (5)

In our case, $d(f^n(x)) = df(f^{n-1}(a)) \circ df^{n-1}(a) = \cdots$ (6)

I'm not sure this will help us. Instead, let's calculate $f^n(x)$ directly, and from that the total derivative. (7)

Let $F\colon S^1 \times D^2 \to S^1 \times D^2$ be the solenoid. Let $x, y \in \mathbb{R}$ and let $\lambda \in (0, \frac{1}{2})$. (8)

Note $F(\phi, x, y) = (2\phi, \lambda x + \frac{1}{2}\cos(2\pi\phi), \lambda y + \sin(2\pi\phi))$. (9)

By writing out the composition, we see that:

$$\begin{aligned} F^n(\phi, x, y)_1 &= 2^n\phi \\ F^n(\phi, x, y)_2 &= \lambda^n x + \frac{1}{2}\lambda^{n-1}\cos(2\pi\phi) + \cdots + \frac{1}{2}\lambda^0\cos(2^{n-1}\pi\phi) \\ &= \lambda^n x + \frac{1}{2}\sum_{i=0}^{n-1}\lambda^i\cos(2^{n-1-i}\pi\phi) \\ F^n(\phi, x, y)_3 &= \lambda^n y + \frac{1}{2}\sum_{i=0}^{n-1}\lambda^i\sin(2^{n-1-i}\pi\phi) \end{aligned} \tag{10}$$

By (9), denoting $\delta_{ij} := \frac{\partial F_i}{\partial z_j}(\phi, x, y)$, we can express $dF^n(\phi, x, y)$ as follows:

$$\begin{aligned}
\delta_{11} &= 2^n \\
\delta_{21} &= -\frac{1}{2}\sum_{i=0}^{n-1} \lambda^i 2^{n-1-i}\pi \sin(2^{n-1-i}\pi\phi) \\
&= -\frac{\pi}{2}\sum_{i=0}^{n-1} \lambda^i 2^{n-1-i} \sin(2^{n-1-i}\pi\phi) \\
\delta_{31} &= \frac{\pi}{2}\sum_{i=0}^{n-1} \lambda^i 2^{n-1-i} \cos(2^{n-1-i}\pi\phi) \\
\delta_{22} &= \lambda^n \\
\delta_{33} &= \lambda^n \\
\delta_{ij} &= 0 \quad \text{otherwise}
\end{aligned} \tag{11}$$

The Lyapunov exponent is defined as

$$\chi(\phi, x, y, v) = \lim_{n\to\infty} \frac{1}{n} \log \|dF^n(\phi, x, y)v\| \tag{12}$$

Note the $\liminf$ in the above definition. (13)

By (10),

$$dF^n(\phi, x, y)v = v_1(2^n + \delta_{21} + \delta_{31}) + \lambda^n(v_2 + v_3) \tag{14}$$

Intuitively, if $v_1 = 0$, then the Lyapunov exponent seems to be $\log(\lambda)$, but if $v_1 \neq 0$, then the first term of $dF^n(\phi, x, y)v$ dominates and the exponent seems to be $\log(2)$, given that δ_{21} and δ_{31} (which depend on n and ϕ) are small enough. (15)

So, we need to argue about the limiting behaviour of these terms. My idea is to bound $\chi(\phi, x, y, v)$ from above and below. We can find an upper bound using the fact that (co)sines are bounded by 1 and the fact that λ is less than $\frac{1}{2}$. (16)

Note, since $\lambda \in (0, \frac{1}{2})$, that $|\delta_{21}| \leq \pi \cdot n \cdot 2^n$ and $|\delta_{31}| \leq \pi \cdot n \cdot 2^n$. (17)

Suppose $v_1 \neq 0$. By (16) and (13), for n sufficiently large,

$$\begin{aligned}
\frac{1}{n}\log\|dF^n(\phi, x, y)v\| &= \frac{1}{2}\cdot\frac{1}{n}\log\left(\|dF^n(\phi, x, y)v\|^2\right) \\
&= \frac{1}{2n}\log\left(v^2(2^n + \delta_{21} + \delta_{31})^2 + \lambda^{2n}(v_2 + v_3)^2\right) \\
&\leq \frac{1}{2n}\log\left((v_1 \cdot 3\pi \cdot n \cdot 2^n)^2 + \lambda^{2n}(v_2 + v_3)^2\right) \\
&\leq \frac{1}{2n}\log\left((v_1 \cdot 4\pi \cdot n \cdot 2^n)^2\right) \\
&= \frac{1}{n}\log(v_1 \cdot 4\pi \cdot n \cdot 2^n) \\
&= \frac{1}{n}\left(\log(v_1 \cdot 4\pi \cdot n) + n\log(2)\right) \\
&\xrightarrow{n\to\infty} \log(2)
\end{aligned} \tag{18}$$

By (17), $\chi(\phi, x, y, v) \leq \log(2)$. (19)

For the lower bound, by (19) and (17),

$$\begin{aligned}\frac{1}{n}\log\|dF^n(\phi,x,y)v\| &= \frac{1}{2n}\log\left(v_1^2(2^n+\delta_{21}+\delta_{31})^2+\lambda^{2n}(v_2+v_3)^2\right)\\ &\geq \frac{1}{2n}\log\left(v_1^2\cdot 2^{2n}\right)\\ &= \frac{1}{n}\log(v_1\cdot 2^n)\\ &= \log(2)+\frac{1}{n}\log(v_1)\xrightarrow{n\to\infty}\log(2)\end{aligned}\tag{20}$$

By (18) and (19), $\chi(\phi,x,y,v)=\log(2)$. (21)

Suppose $v_1=0$. By (21) and (13),

$$\begin{aligned}\frac{1}{n}\log\|dF^n(\phi,x,y)v\| &= \frac{1}{n}\log(\lambda^n(v_2+v_3))\\ &= \log(\lambda)+\frac{1}{n}\log(v_2+v_3)\\ &\xrightarrow{n\to\infty}\log(\lambda)\end{aligned}\tag{22}$$

By (21), $\chi(\phi,x,y,v)=\log(\lambda)$. (23)

By (22) and (20), the Lyapunov exponents are $\log(2)$ and $\log(\lambda)$. $\square$ (24)

Exercise 2.1.3.

Let $f : X \to X$ be a topological dynamical system. (1)

Show that $\mathcal{R}(f) \subseteq \mathrm{NW}(f)$. (2)

Proof + reasoning:

Let $x \in \mathcal{R}(f)$. (3)

Let U be a neighborhood of x, and V an open set such that $V \subseteq U$ and $x \in V$. (4)

We need to show that there exists an $n \geq 1$ such that $f^n(U) \cap U \neq \emptyset$. (5)

By (3) and (4), there exists a recurrent point y in U. (6)

By (6), there exists an increasing sequence (n_k) such that

$$f^{n_k}(y) \to y \quad \text{and} \quad n_k \to \infty. \tag{7}$$

Intuitively, the sequence of sets $f^{n_k}(U)$ should eventually intersect with U, giving our required result. (8)

A possible problem in the proof so far is that y may not be in the interior of U, so x is not necessarily in a neighborhood of y, so it is possible that $f^{n_k}(U)$ comes arbitrarily close to y but never intersects with U. (9)

I think we can avoid this problem by making a stronger statement than (6), using V instead of U: (10)

By (3) and (4), there exists a recurrent point z in V. (11)

By (11), there exists an increasing sequence (m_k) such that

$$f^{m_k}(z) \to z \quad \text{and} \quad m_k \to \infty. \tag{12}$$

Since V is a neighborhood of z, by (12) there exists an $M \geq 1$ such that $\forall i \geq M$, $f^{m_i}(z) \in V$, so $f^{m_M}(z) \in U$, hence $f^{m_M}(U) \cap U \neq \emptyset$. (13)

By (13), $\mathcal{R}(f) \subseteq \mathrm{NW}(f)$. (14)

Exercise 2.2.3.

Is the product of two topologically transitive systems topologically transitive? (1)

Is a factor of a topologically transitive system topologically transitive? (2)

Proof + reasoning:

I think the statement is false. Let's construct a counterexample, by starting with the simplest non-trivial product of topologically transitive systems, and iterating from there. (3)

Let R_α be the circle translation, where α is irrational. (4)

It is known that R_α is topologically transitive. (5)

$R_\alpha \times R_\alpha$ already seems to form a counterexample, since $R_\alpha \times R_\alpha$ preserves the distance between the components of points in $S^1 \times S^1$. (6)

Let $(a, b) \in S^1 \times S^1$. (7)

If $a \geq b$, then the orbit of (a, b) under $R_\alpha \times R_\alpha$ is contained in $l_1 \cup l_2$ where

$$\begin{aligned} l_1 &= \{t(a-b, 0) + (1-t)(1, b-a+1) : t \in [0,1]\}\,, \\ l_2 &= \{t(1, b-a+1) + (1-t)(1, b-a+1) : t \in [0,1]\}\,. \end{aligned} \tag{8}$$

If $b \geq a$, then the same holds with

$$\begin{aligned} l_1 &= \{t(0, a-b+1) + (1-t)(b-a, 1) : t \in [0,1]\}\,, \\ l_2 &= \{t(b-a, 0) + (1-t)(1, a-b+1) : t \in [0,1]\}\,. \end{aligned} \tag{9}$$

In both cases, l_1 and l_2 are lines contained in $[0, 1) \times [0, 1)$. Since these lines are clearly not dense in $S^1 \times S^1$, the forward orbit of (a, b) is not dense in $S^1 \times S^1$. Hence, $R_\alpha \times R_\alpha$ is not topologically transitive. (10)

Suppose $f : X \to X$ and $g : Y \to Y$ are topological dynamical systems, that π is a topological semiconjugacy from f to g, and that f is topologically transitive with point $x \in X$ with dense forward orbit. (11)

We want to show that the forward orbit of $\pi(x)$ is dense. (12)

Let $U \subseteq Y$ be open. Since π is continuous, $\pi^{-1}(U)$ is open, so by (11), there exists a $k \in \mathbb{N}$ such that $f^k(x) \in \pi^{-1}(U)$. (13)

By (11), $\pi \circ f^k(x) = g^k(\pi(x))$. (14)

By (11) and (13), $g^k(\pi(x)) \in U$, so $\pi(x)$ is dense, hence a factor of a topologically transitive system is topologically transitive. (15)

Exercise 2.3.3.

Show that a factor of a topologically mixing system is also topologically mixing. (1)

Proof

Let $f : X \to X$ and $g : Y \to Y$ be topological dynamical systems, and π a topological semiconjugacy from f to g. (2)

Let U and V be nonempty open sets in Y. (3)

Since π is surjective and continuous, $\pi^{-1}(U)$ and $\pi^{-1}(V)$ are nonempty and open. (4)

By (1) and (3), there exists an $N \in \mathbf{N}$ such that for all $n \geq N$, $f^n(\pi^{-1}(U)) \cap \pi^{-1}(V) \neq \varnothing$. (5)

By (1),

$$\begin{aligned} \pi\left(f^n(\pi^{-1}(U)) \cap \pi^{-1}(V)\right) &\subseteq \pi(f^n(\pi^{-1}(U))) \cap \pi(\pi^{-1}(V)) \\ &= g^n(\pi(\pi^{-1}(U))) \cap \pi(\pi^{-1}(V)) \\ &= g^n(U) \cap V. \end{aligned} \tag{6}$$

By (5) and (4), $g^n(U) \cap V \neq \varnothing$, so g is topologically mixing, hence a factor of a topologically mixing system is topologically mixing. (7)

Exercise 2.5.3.

Let $\{a_n\}$ be a subadditive sequence of non-negative real numbers, i.e. (1)

$0 \leq a_{m+n} \leq a_m + a_n$ for all $m, n \geq 0$. (2)

Show that $\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 0} \frac{a_n}{n}$. (3)

Proof + reasoning:

We need to show $\lim_{n\to\infty} \frac{a_n}{n}$ is the infimum, which we can do by checking that it is a lower bound of $\left\{\frac{a_n}{n}\right\}$, and that it is greater than or equal to any lower bound of $\left\{\frac{a_n}{n}\right\}$. (4)

Let $k \in \mathbb{N}_+$. (5)

We need to show $\lim_{n\to\infty} \frac{a_n}{n} \leq \frac{a_k}{k}$. To do that, we could find $f(n) \to 0$ such that, for n sufficiently large,

$$\frac{a_n}{n} \leq \frac{a_k}{k} + f(n).$$

What natural choice for n can we make? (6)

Let $n \geq k$. (7)

To prove (6), we will need to show that $\frac{a_n}{n} - \frac{a_k}{k} \to 0$ as $n \to \infty$. This should follow in some way from the nonnegativity and the subadditivity of (a_i). (8)

Observe that for any natural number m, the value a_{mk} is bounded by ma_k, therefore,

$$\frac{a_{mk}}{mk} \leq \frac{ma_k}{mk} = \frac{a_k}{k}.$$

In other words, for the subsequence of a_i where i is a multiple of k we have the required convergence. And, again by subadditivity, any other element is at most ka_1 away from this subsequence. (9)

By (7), $n = mk + m'$, where $m \in \mathbb{N}$ and $m' < k$. (10)

By (10), and the subadditivity of (a_n),

$$\begin{aligned}\frac{a_n}{n} - \frac{a_k}{k} &= \frac{a_{mk+m'}}{n} - \frac{a_k}{k}\\ &\leq \frac{a_{mk} + a_{m'}}{n} - \frac{a_k}{k}\\ &\leq \frac{ma_k}{mk+m'} + \frac{ka_1}{n} - \frac{a_k}{k}\\ &\xrightarrow{n\to\infty} 0\end{aligned} \tag{11}$$

Hence, $\lim_{n\to\infty} \frac{a_n}{n}$ is a lower bound for $\left\{\frac{a_n}{n} : n \geq 1\right\}$. (12)

Additionally, if $C \leq \frac{a_m}{m}$ for all $m \in \mathbb{N}$, then clearly $C \leq \lim_{n\to\infty} \frac{a_n}{n}$, so by (12),

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 0} \frac{a_n}{n}. \tag{13}$$

Exercise 2.7.3.

Give a non-trivial example of a homeomorphism f of a compact metric space (X, d) such that $d(f^n(x), f^n(y)) \to 0$ as $n \to \infty$ for every pair $x, y \in X$. (1)

Proof + reasoning:

Let's recall simple examples of continuous maps of compact metric spaces, and see if they satisfy (1). (2)

The translation R_α on S^1 is not valid, as it preserves distances. (3)

The map $h : x \mapsto \frac{1}{2}x$ on S^1 has the property that $d(h^n(x), h^n(y)) \to 0$ as $n \to \infty$ for every pair $x, y \in X$, but it is not surjective. (4)

To solve this issue, we can define a map that is the piecewise combination of a contraction (such as f) on one half of the circle and an expansion on the other half of the circle. (5)

Define $f : S^1 \to S^1$ by

$$f(x) = \begin{cases} \frac{1}{2}x & \text{if } x \in [0, \frac{1}{2}) \\ \frac{3}{2}x - \frac{1}{2} & \text{if } x \in [\frac{1}{2}, 1) \end{cases} \tag{6}$$

Clearly, f is a homeomorphism such that $d(f^n(x), f^n(y)) \to 0$ as $n \to \infty$ for all pairs x, y. (7)

Exercise 2.8.3.

Prove the following generalization of Proposition 2.1.2. If a commutative group G acts by homeomorphisms on a compact metric space X, then there is a non-empty, closed G-invariant subset X' on which G acts minimally. (1)

Proof + reasoning:

Let's try to adapt the proof of Proposition 2.1.2 to the more general case. (2)

There are four theorems stated in section 2.8, but they don't seem applicable in this exercise. (3)

Let $\mathcal{C}$ be the collection of non-empty, closed G-invariant subsets of X, with the partial ordering given by inclusion. (4)

Since $X \in \mathcal{C}$, $\mathcal{C}$ is not empty. (5)

Suppose $\mathcal{K} \subseteq \mathcal{C}$ is a totally ordered subset. Then, any finite intersection of elements of $\mathcal{K}$ is nonempty, so by the finite intersection property for compact sets, $\bigcap_{K\in\mathcal{K}} K \neq \emptyset$. Thus, by Zorn's lemma, $\mathcal{C}$ contains a minimal element M. (6)

So far, we have followed the proof of 2.1.2 almost exactly. How can we conclude that G acts minimally on M? (7)

Intuitively speaking, we have very little constructive information about M, so proof by contradiction seems like a good strategy. (8)

Suppose that G does not act minimally on M. (9)

Then, there exists a point $b \in M$ and a nonempty open set $C \subseteq M$ such that $Gb \cap C = \emptyset$. (10)

To conclude the proof, we should find a contradiction, given (9) and (5). We want to find some set $Y \subseteq M$ that is closed, nonempty, and G-invariant. Let's start with some possible natural choices for Y, and work from there. (11)

Note C is not closed, $\mathrm{cl}(C)$ is not necessarily G-invariant, and $M \setminus C$ is also not necessarily G-invariant. On the other hand, $M \setminus GC$ seems like a good candidate. (12)

Since $GC = \bigcup_{g\in G} gC$, and each $g \in G$ is a homeomorphism, GC is open. (13)

So, since M is closed, $M \setminus GC$ is closed. (14)

Since $b \in M$ and $Gb \cap C = \emptyset$, $b \in M \setminus GC$, so $M \setminus GC$ is nonempty. (15)

If $m \in M \setminus GC$ and $g \in G$, then $m \neq g^{-1}c$, so $gm \neq C$, so since M is G-invariant, $gm \in M \setminus GC$, so $M \setminus GC$ is G-invariant. (16)

By (12)–(14), $M \setminus GC$ is a closed, nonempty, G-invariant proper subset of M, which contradicts (5), so (8) is false. Hence, G acts minimally on M. (17)

Exercise 3.1.3.

Use a higher block presentation to prove that for any block code $c : X \to Y$ there is a subshift Z and an isomorphism $f : Z \to X$ such that $c \circ f : Z \to Y$ is a $(0,0)$-block code. (1)

Proof + reasoning:

Let $c : X \to Y$ be a block code, with corresponding function $\alpha : W_{a+b+1} \to \mathcal{A}_m$. (2)

Clearly, any higher block presentation f of X gives an isomorphism $f : X \to \mathrm{im}(f)$, but it is not clear which one will make $f \circ c$ a $(0,0)$-block code. Intuitively, the most natural choice is the presentation of X in which the blocks are aligned with those given by c. Let's check if this presentation has the required properties. (3)

Letting $k = a + b + 1$ and $l = b$, the higher block presentation d of X can be written as

$$d(x)_i = x_{i-a} \dots x_{i+b}, \quad i \in \mathbb{Z} \tag{4}$$

Since $\mathrm{im}(d) \subseteq \Sigma_{W_{a+b+1}(X)}$ we have $W_1(\mathrm{im}(d)) \subseteq W_{a+b+1}(X)$. (5)

If $\omega \in W_{a+b+1}(X)$, then for some sequence $x \in X$ and $i \in \mathbb{Z}$,

$$\omega = x_{i-a} \dots x_{i+b}, \quad \text{so} \quad d(x)_i = \omega,$$

so $\omega \in W_1(\mathrm{im}(d))$. (6)

By (5) and (4), $W_1(\mathrm{im}(d)) = W_{a+b+1}(X)$. (7)

By Exercise 3.1.2, d is an isomorphism onto its image. Let $d^{-1} : \mathrm{im}(d) \to X$ be its inverse. (8)

If $z \in \mathrm{im}(d)$ and $i \in \mathbb{Z}$, then there exists a unique x such that $d(x) = z$, so

$$\begin{aligned}(c \circ d^{-1})(z)_i &= c(d^{-1}(z))_i \\ &= c(x)_i \\ &= \alpha(x_{i-a} \dots x_{i+b}) \\ &= \alpha(z_i)\end{aligned} \tag{9}$$

By (7) and (5), $c \circ d^{-1} : \mathrm{im}(d) \to Y$ is a $(0,0)$-block code and, by (6), d^{-1} is an isomorphism. (10)

Exercise 3.2.3.

Show that every edge shift is an SFT. (1)

If Σ_B^e is an edge shift with graph Γ_B, then Σ_B^e is precisely the set of sequences that do not contain the words $e'e$ of length 2 in which the target of e is not equal to the source of e'. (2)

Since this collection of words is finite, Σ_B^e is an SFT. (3)

Exercise 4.2.3

Prove that if T is a measure-preserving transformation, then so are the induced transformations. (1)

Proof + reasoning:

Let $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ be a measure-preserving transformation. (2)

The primitive transformation is similar to a suspension with ceiling f. (3)

What is the natural measure on $(X_f, \mathcal{F})$? The only obvious one is the product of μ and the counting measure on $\mathbb{N}$. Although not explicitly stated, this is exactly the measure μ_f (4)

Let's start with proving that the derivative transformation is measure preserving. First, let's check that it is measurable. (5)

What is the natural σ-algebra on A? Clearly, it should be the 'subspace' sigma algebra, also referred to as the trace σ-algebra. (6)

Let $\mathcal{E}$ be the trace σ-algebra with respect to $A \in \mathcal{A}$. (7)

Let $B \in \mathcal{E}$. (8)

By definition, $B = C \cap A$, $C \in \mathcal{A}$. (9)

We want to show that $T_A^{-1}(B) \in \mathcal{E}$, i.e. that $T_A^{-1}(B) = B' \cap A$ for some $B' \in \mathcal{E}$. (10)

A standard way to show that a set such as $T_A^{-1}(B)$ is measurable is to equate it to a countable combination of measurable sets. (11)

$$T_A^{-1}(B) = \bigcup_{k \geq 1} (T^{-k} \cap A)? \tag{12}$$

Statement (12) seems false, in that $\bigcup_{k \geq 1} (T^{-k} \cap A)$ contains points that aren't in $T_A^{-1}(A)$, i.e. given $y \in A$, if $x \in T^{-1}(y)$, and $\exists z \in A$ s.t. $T(z) = x$, then $z \in \bigcup_{k=1}^{\infty} T^{-k}(y)$ but $z \notin T_A^{-1}(y)$. (13)

Let's try a different idea. Intuitively, the set $T^{-1}(B)$ can be partitioned into $T^{-1}(B) \cap A$, which clearly is a subset of $T_A^{-1}(B)$, and the 'remainder', $T^{-1}(B) \setminus A$. The remainder may still contain points x of which the inverse image 'eventually' intersects with A, meaning that there exists a $k \in \mathbf{N}$ with $T^{-k}(x) \in A$. By repeatedly taking the inverse image of the remainder and intersecting with A, we should obtain all points in $T_A^{-1}(B)$. Let's formalize this. (14)

Let (R_n) and (D_n) be sequences of sets defined inductively by letting

$$\begin{aligned} R_0 &= B, & D_0 &= \emptyset \\ R_1 &= T^{-1}(B) \setminus A, & D_1 &= T^{-1}(B) \cap A \\ R_{n+1} &= T^{-1}(R_n) \setminus A, & D_{n+1} &= T^{-1}(R_n) \cap A \quad \forall n \geq 2 \end{aligned} \tag{15}$$

Let $D := \bigcup_{n \geq 1} D_n$ (16)

Conjecture 3. $D = T_A^{-1}(B)$

Proof. Let $n \geq 2$

$$\begin{aligned} R_n &= T^{-1}(R_{n-1}) \setminus A \\ &= T^{-1}(R_{n-1}) \cap A^c \\ &= T^{-1}(T^{-1}(R_{n-2}) \cap A^c) \cap A^c \\ &= (T^{-2}(R_{n-2}) \cap T^{-1}(A^c)) \cap A^c \\ &= (T^{-n}(B) \cap \cdots \cap T^{-1}(A^c)) \cap A^c \\ &= T^{-n}(B) \cap \left(\bigcap_{i=0}^{n-1} T^{-i}(A^c) \right). \end{aligned}$$

This gives

$$D_n = T^{-n}(B) \cap \left(\bigcap_{i=1}^{n-1} T^{-i}(A^c) \right) \cap A.$$

By definition, $T_A^{-1}(B)$ is the set of points $y \in A$ such that $T(y) \in B$ or such that there exists a $k \in \mathbb{N}$ with $k \geq 2$ such that $T^k(y) \in B$ and $T^i(y) \notin A$ for all $i \in \{1, \ldots, k-1\}$.

From the above, , it follows that $D = T_A^{-1}(B)$. □

T is $\mathcal{A}$-measurable, so from $D_n = T^{-n}(B) \cap \left(\bigcap_{i=1}^{n-1} T^{-i}(A^c) \right) \cap A$ it follows that $D_n \in \mathcal{E}$. Since D is a countable union of such D_n, $D \in \mathcal{E}$, so by conjecture 3, $T_A^{-1}(B) \in \mathcal{E}$, so T_A is $\mathcal{E}$-measurable. (17)

Now we need to show that T_A is measure-preserving. I think the sets D_n are disjoint. Let's prove this using contradiction. (18)

Let $i, j \in \mathbb{N}$ with $i > j \geq 1$. (19)

Suppose $D_j \cap D_i \neq \emptyset$. (20)

By (20), there exists $x \in D_j \cap D_i$. By the fact that $D_n = T^{-n}(B) \cap \left(\bigcap_{i=1}^{n-1} T^{-i}(A^c) \right) \cap A$, it follows that $T^j(x) \in B$. (21)

Since $j < i$, $T^j(x) \in A^c$, but this contradicts (21), so D_j and D_i are disjoint. (22)

We haven't used any results from the given section. The only theorem mentioned in section 4.2 is the Poincaré recurrence theorem, which essentially states that almost all points in B eventually return to B. In our case, it seems to imply that $\mu(B)$ is a lower bound of $\mu(D)$. (23)

What can I say about the measure of the sequence D_n and R_n? Intuitively, at each step n, the measure of R_n is preserved, but divided among D_{n+1} and R_{n+1}. This, together with the disjointness of D_i, seems to imply that the measure of all D_i up to D_n together with the measure of R_n is constant over time, and equal to $\mu(R_0) = \mu(B)$. That could give us an upper bound for $\mu(D)$, which together with (23) might allow us to conclude. (24)

From (15), it is clear that $\forall n \in \mathbb{N}$, $D_n \cap R_n = \emptyset$ (25)

$$\mu(R_{n+1}) + \mu(D_{n+1}) = \mu(R_n) \quad \forall n \in \mathbb{N} \tag{26}$$

From (37), $\mu(R_n)$ is decreasing. (27)

By (22),

$$\mu(D) = \sum_{n \geq 1} \mu(D_n) \tag{28}$$

From (37),

$$\sum_{1 \leq i \leq n} \mu(D_n) = \mu(B) - \mu(R_n) \quad \forall n \in \mathbb{N} \tag{29}$$

By (28), (29) and (38),

$$\mu(D) = \lim_{n \to \infty} \sum_{1 \leq i \leq n} \mu(D_n) = \mu(B) - \lim_{n \to \infty} \mu(R_n) \tag{30}$$

By the Poincaré recurrence theorem,

$$\mu(T_A^{-1}(B)) \geq \mu(B) \tag{31}$$

By conjecture 3 and (31), $\mu(D) \geq \mu(B)$. (32)

By (45) and (41),

$$\mu(B) - \lim_{n \to \infty} \mu(R_n) = \mu(D) \geq \mu(B) \tag{33}$$

From (33),

$$\lim_{n \to \infty} \mu(R_n) = 0 \tag{34}$$

From (34), (41) and conjecture 3,

$$\mu(T_A^{-1}(B)) = \mu(D) = \mu(B) \tag{35}$$

By (35), T_A is measure-preserving. (36)

Next, we want to show that the primitive transformation is measure-preserving. (37)

Let $T_f : X_f \to X_f$ be the primitive transformation, where $f : X \to \mathbb{N}$ is measurable. (38)

By basic measure theory, it suffices to show, to prove T_f is measure-preserving, that T_f preserves the measure of all elements of a generating set of the σ-algebra. (39)

Let $A \in \mathcal{A}$ and $k \in \mathbb{N}$. (40)

Note, $(A \times \{k\}) \cap X_f = (A \cap C_k) \times \{k\}$ where $C_k = f^{-1}(\{n \in \mathbb{N} : n \geq k\})$ (41)

$X_f = \{(x, k) : x \in X, 1 \leq k \leq f(x)\} \subseteq X \times \mathbb{N}$ (42)

Suppose $k > 1$. By (41):

$$T_f^{-1}((A \times \{k\}) \cap X_f) = T_f^{-1}((A \cap C_k) \times \{k\}) = (A \cap C_k) \times \{k - 1\} \tag{43}$$

Suppose $k = 1$ (44)

$$T_f^{-1}(A \times \{k\}) = \bigcup_{i \geq 1} (T^{-1}(f^{-1}(i) \cap A) \times \{i\}) \tag{45}$$

From (43) and (45), $T_f^{-1}(A \times \{k\}) \cap X_f \in \mathcal{U}_f \quad \forall k$ (46)

Now we show that T_f preserves the measure: (47)

If $k > 1$, then by (43) and (41),

$$\begin{aligned}\mu_f(T_f^{-1}((A \times \{k\}) \cap X_f)) &= \mu_f((A \cap C_k) \times \{k-1\}) \\ &= \mu_f((A \cap C_k) \times \{k\} \cap X_f)\end{aligned} \tag{48}$$

If $k = 1$, then by (45) and T being measure-preserving,

$$\begin{aligned}\mu_f(T_f^{-1}(A \times \{1\})) &= \mu_f\left(\bigcup_{i\geq 1}(T^{-1}(f^{-1}(i) \cap A) \times \{i\})\right) \\ &= \sum_{i\geq 1}\mu_f(T^{-1}(f^{-1}(i) \cap A) \times \{i\}) \\ &= \sum_{i\geq 1}\mu(T^{-1}(f^{-1}(i) \cap A)) \\ &= \mu\left(\bigcup_{i\geq 1}T^{-1}(f^{-1}(i) \cap A)\right) \\ &= \mu(T^{-1}(f^{-1}(\mathbb{N}) \cap A)) \\ &= \mu(T^{-1}(A)) \\ &= \mu(A) \\ &= \mu_f(A \times \{1\})\end{aligned} \tag{49}$$

By (48) and (49), the primitive transformation is measure-preserving. (50)

Exercise 4.3.3.

A measure-preserving transformation or flow T of a probability space $(X,\mathcal{U},\mu)$ is called *(strong) mixing* if

$$\lim_{t\to\infty}\mu(T^t(A)\cap B)=\mu(A)\cdot\mu(B)$$

for any two measurable sets $A,B\in\mathcal{U}$. (1)

Equivalently, T is mixing if

$$\lim_{t\to\infty}\int_X f(T^t(x))g(x)\,d\mu=\int_X f\,d\mu\int_X g\,d\mu$$

for any two bounded measurable functions. (2)

Transformation T is called *weak mixing* if $\forall A,B\in\mathcal{U}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(T^{-i}(A)\cap B)-\mu(A)\mu(B)\right|=0. \tag{3}$$

Equivalently, T is weak mixing if for all bounded measurable functions,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\int_X f(T^i(x))g(x)\,d\mu-\int_X f\,d\mu\int_X g\,d\mu\right|=0. \tag{4}$$

Flow T is called weak mixing if $\forall A,B\in\mathcal{U}$,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\left|\mu(T^{-s}(A)\cap B)-\mu(A)\mu(B)\right|\,ds=0. \tag{5}$$

Equivalently, T is weak mixing if for all bounded measurable functions,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\left|\int_X f(T^s(x))g(x)\,d\mu-\int_X f\,d\mu\int_X g\,d\mu\right|ds=0. \tag{6}$$

Show that the two definitions of strong and weak mixing given in terms of sets and bounded measurable functions are equivalent. (7)

Proof + reasoning:

Let's start with the most complicated implication, namely (5) $\Rightarrow$ (6). (8)

Let T be a measure-preserving flow on $(X,\mathcal{U},\mu)$. (9)

Assume (5) holds. (10)

Let's make some observations first. The left-hand side of (5) contains only one integral, unlike that of (5). My guess is that we can relate these via "standard machinery", which means to prove the equivalence for f,g simple, and then use the fact that f and g are both limits of simple functions, together with suitable convergence theorems. (11)

Suppose f and g are simple, with

$$f=\sum_{i\le n}\mathbf{1}_{A_i}a_i,\quad g=\sum_{j\le n}\mathbf{1}_{A_j}b_j. \tag{12}$$

Define

$$M := \lim_{t\to\infty} \frac{1}{t}\int_0^t \left|\int_X f(T^s(x))g(x)\,d\mu - \int_X f\,d\mu \int_X g\,d\mu\right| ds. \tag{13}$$

Then

$$M = \lim_{t\to\infty} \frac{1}{t}\int_0^t \left|\int_X f(T^s(x))g(x)\,d\mu - \sum_{i,j\le n} \mu(A_i)\mu(A_j)a_ib_j\right| ds. \tag{14}$$

How can we simplify $\int_X f(T^s(x))g(x)\,d\mu$? I think we can write the composition of a simple function after an arbitrary function as a simple function: (15)

Let $h : \mathbb{R} \to \mathbb{R}$ be an arbitrary function. For all $x \in \mathbb{R}$, we have that $f(h(x)) = a_i$ if $h(x) \in A_i$. Hence, for all $x \in \mathbb{R}$,

$$f(h(x)) = \sum_{i\le n} a_i \mathbf{1}_{h_{-1}(A_i)}(x). \tag{16}$$

By (16),

$$\begin{aligned}\int_X f(T^s(x))g(x)\,d\mu &= \int_X \left(\sum_{i\le n} \mathbf{1}_{T^{-s}(A_i)}a_i\right)\left(\sum_{j\le n} \mathbf{1}_{A_j}b_j\right) d\mu \\ &= \sum_{i,j\le n} \mu(T^{-s}(A_i)\cap A_j)a_ib_j.\end{aligned} \tag{17}$$

Clearly,

$$\int_X f\,d\mu \int_X g\,d\mu = \left(\sum_{i\le n} \mu(A_i)a_i\right)\left(\sum_{j\le n} \mu(A_j)b_j\right) = \sum_{i,j\le n} \mu(A_i)\mu(A_j)a_ib_j. \tag{18}$$

By (17) and (18),

$$M = \lim_{t\to\infty} \frac{1}{t}\int_0^t \left|\sum_{i,j\le n} a_ib_j \left(\mu(T^{-s}(A_i)\cap A_j) - \mu(A_i)\mu(A_j)\right)\right| ds. \tag{19}$$

We would like to interchange the summation and the absolute value operation, but the absolute value operation is not linear, except if each term is positive. Perhaps, if we prove the statement for positive simple functions first, we can conclude for all bounded simple functions? (20)

That seems unnecessary. Crucially, M is nonnegative so we only need to prove 0 is an upper bound and we can still use the subadditive property of the absolute value, i.e., $|a+b| \le |a| + |b|$: (21)

By (19),

$$\begin{aligned}
M &\leq \lim_{t\to\infty} \frac{1}{t} \int_0^t \sum_{i,j\leq n} |a_i b_j| \left|\mu(T^{-s}(A_i)\cap A_j) - \mu(A_i)\mu(A_j)\right| ds \\
&= \sum_{i,j\leq n} |a_i b_j| \lim_{t\to\infty} \frac{1}{t} \int_0^t \left|\mu(T^{-s}(A_i)\cap A_j) - \mu(A_i)\mu(A_j)\right| ds \\
&= 0.
\end{aligned} \tag{22}$$

Now we can prove the equivalence for general f, g. (23)

Assume that f and g are measurable and bounded by some $C > 0$. (24)

By (36), f and g are the uniform limits of sequences (f_n) and (g_n) respectively, where f_n and g_n are simple functions that are bounded by C. (25)

By the dominated convergence theorem,

$$\begin{aligned}
M &= \lim_{t\to\infty} \frac{1}{t} \int_0^t \left| \int_X \lim_{n\to\infty} f_n(T^s(x)) \lim_{n\to\infty} g_n(x)\, d\mu - \int_X \lim f_n\, d\mu \int_X \lim g_n\, d\mu \right| ds \\
&= \lim_{t\to\infty} \frac{1}{t} \int_0^t \lim_{n\to\infty} \left| \int_X f_n(T^s(x)) g_n(x)\, d\mu - \int_X f_n\, d\mu \int_X g_n\, d\mu \right| ds
\end{aligned} \tag{26}$$

Can we place the limit outside? (27)

Note,

$$\left| \int_X f_n(T^s(x)) g_n(x)\, d\mu - \int_X f_n\, d\mu \int_X g_n\, d\mu \right| \leq 2C\mu(X). \tag{28}$$

By (26), by the fact that the absolute value is continuous, and by (28) together with the dominated convergence theorem,

$$\begin{aligned}
M &= \lim_{t\to\infty} \frac{1}{t} \int_0^t \lim_{n\to\infty} \left| \int_X f_n(T^s(x)) g_n(x)\, d\mu - \int_X f_n\, d\mu \int_X g_n\, d\mu \right| ds \\
&= \lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{t} \int_0^t \left| \int_X f_n(T^s(x)) g_n(x)\, d\mu - \int_X f_n\, d\mu \int_X g_n\, d\mu \right| ds
\end{aligned} \tag{29}$$

Can we switch the limits? Is there any general result that can help? Yes, we might be able to use the Moore–Osgood theorem. (30)

This theorem states that the two limits are interchangable under certain conditions, including the condition that

$$h_n(t) := \frac{1}{t} \int_0^t \left| \int_X f_n(T^s(x)) g_n(x)\, d\mu - \int_X f\, d\mu \int_X g\, d\mu \right| ds$$

has the property that there exists a $\delta > 0$ such that $h_n(t) : (0, \delta) \to \mathbb{R}$ converges uniformly to some limit as $n \to \infty$. (31)

Let $\mathcal{T} = (0, \delta)$, where $\delta > 0$. (32)

Let $\Delta_1 > 0$. (33)

There exists a $k \in \mathbb{N}$ such that for all $n \geq k$ and all $x \in X$, $f_n(T^s(x))g_n(x) \in B(f(T^s(x))g(x), \Delta_1)$, where $B(x, \epsilon)$ denotes a ball around x of radius ϵ, so

$$\int_X f_n(T^s(x))g_n(x)\,d\mu \in B\left(\int_X f(T^s(x))g(x)\,d\mu, \Delta_1\mu(X)\right) \tag{34}$$

Let $\Delta_2 \geq 0$. Then there exists an m such that for $n \geq m$,

$$\int_X f_n\,d\mu \in B\left(\int_X f\,d\mu, \Delta_2\mu(X)\right) \quad \text{and} \quad \int_X g_n\,d\mu \in B\left(\int_X g\,d\mu, \Delta_2\mu(X)\right). \tag{35}$$

By (34) and (35), for all $n \geq \max(m, k)$,

$$\begin{aligned}&\frac{1}{t}\int_0^t \left|\int_X f_n(T^s(x))g_n(x)\,d\mu - \int_X f\,d\mu \int_X g\,d\mu\right| ds \\ &\in B\Big(\frac{1}{t}\int_0^t \left|\int_X f(T^s(x))g(x)\,d\mu - \int_X f\,d\mu \int_X g\,d\mu\right| ds, \Delta_1\mu(X) + 2\Delta_2\mu(X)\Big)\end{aligned} \tag{36}$$

Since Δ_1 and Δ_2 were arbitrary, and (36) does not depend on t,

$$\|h_n(t) - h(t)\|_{\mathcal{T}} \to 0. \tag{37}$$

Therefore, by the Moore–Osgood theorem,

$$M = \lim_{n\to\infty}\lim_{t\to\infty}\frac{1}{t}\int_0^t \left|\int_X f_n(T^s(x))g_n(x)\,d\mu - \int_X f_n\,d\mu \int_X g_n\,d\mu\right| ds = 0. \tag{38}$$

Clearly, (6) implies (5), so (5) and (6) are equivalent. (39)

We will skip the proof of (3) $\Leftrightarrow$ (4), since it is likely very similar to the one for (5) $\Leftrightarrow$ (6). (40)

Next, suppose

$$\lim_{t\to\infty}\mu(T^{-t}(A) \cap B) = \mu(A)\cdot\mu(B)$$

for any two measurable sets $A, B \in \mathcal{U}$. (41)

By (41) and by dominated convergence,

$$\begin{aligned}\lim_{t\to\infty}\int_X f_n(T^{-t}(x))g_n(x)\,d\mu &= \lim_{t\to\infty}\int_X \left(\sum_{i=1}^n \mathbf{1}_{T^{-t}(A_i)}a_i^n\right)\left(\sum_{j=1}^n \mathbf{1}_{A_j}b_j^n\right) d\mu \\ &= \lim_{t\to\infty}\int_X \left(\sum_{i,j\leq n}\mathbf{1}_{T^{-t}(A_i)\cap A_j}(x)a_i^n b_j^n\right) d\mu \\ &= \sum_{i,j\leq n}\lim_{t\to\infty}\int_X \mathbf{1}_{T^{-t}(A_i)\cap A_j}(x)a_i^n b_j^n\,d\mu \\ &= \sum_{i,j\leq n}\lim_{t\to\infty}\mu(T^{-t}(A_i)\cap A_j)a_i^n b_j^n \\ &= \sum_{i,j\leq n}\mu(A_i)\mu(A_j)a_i^n b_j^n.\end{aligned} \tag{42}$$

So,

$$\int_X f\,d\mu \int_X g\,d\mu. \tag{43}$$

Using (25), (43) and dominated convergence, it follows from an argument similar to the one used to derive (38) that

$$\begin{aligned}
M_s(f,g) &:= \lim_{t\to\infty} \int_X f(T^t(x))g(x)\,d\mu \\
&= \lim_{t\to\infty} \int_X \lim_{n\to\infty} f_n(T^t(x)) \lim_{n\to\infty} g_n(x)\,d\mu \\
&= \lim_{t\to\infty} \lim_{n\to\infty} \int_X f_n(T^t(x))g_n(x)\,d\mu \\
&= \lim_{n\to\infty} \lim_{t\to\infty} \int_X f_n(T^t(x))g_n(x)\,d\mu \\
&= \lim_{n\to\infty} \int_X f_n\,d\mu \int_X g_n\,d\mu \\
&= \int_X f\,d\mu \int_X g\,d\mu.
\end{aligned} \tag{44}$$

In other words, statement (1) implies statement (2). Since the converse is trivial, (1) and (2) are equivalent. (45)