Lecture #3

In the last lecture, we saw that hyperbolic toral automorphisms are mixing:

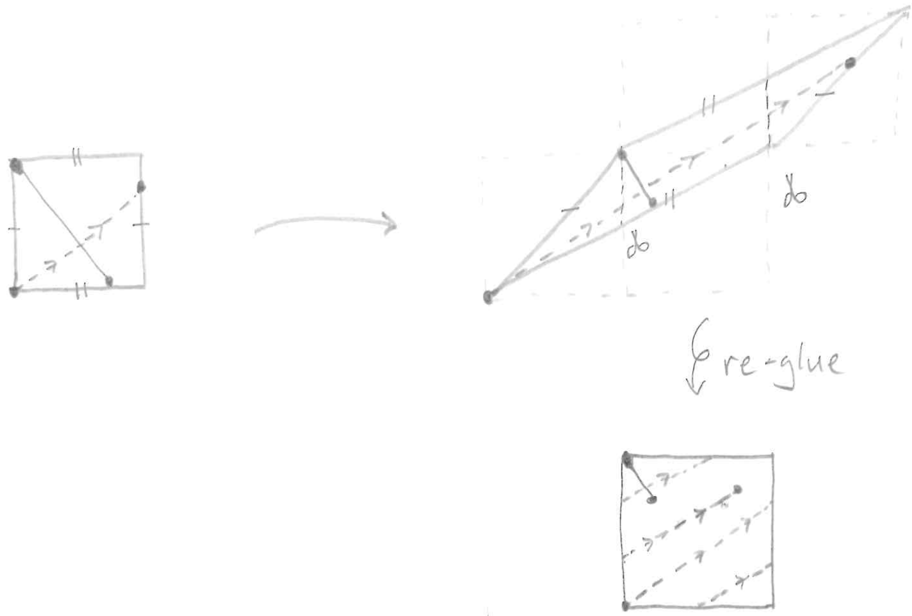
<u>THM A</u> Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ and let A be an $n \times n$ matrix with integer coefficients, determinant 1, and no eigenvalues on the unit circle. Then, the hyperbolic toral automorphism $A: \mathbb{T}^n \longrightarrow \mathbb{T}^n$ is mixing.

This lecture will be devoted to re-proving this result using a new argument in lieu of the Fourier analysis used in the previous lecture. This new argument will work in much more general contexts, especially in contexts where Fourier (or Harmonic) analysis is not so explicit or easy to work with.

§ Coudène-Hopf Argument

The argument that we are going to develop to re-prove that hyperbolic toral automorphisms are mixing derives from the classical Hopf argument used to prove ergodicity of dynamical systems that have sufficiently many expanding and contracting "directions" (or perhaps better said, no directions which are not either expanding or contracting).

Let's look at a concrete example to make this intuitive. Consider the map $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \curvearrowright \mathbb{T}^2$:

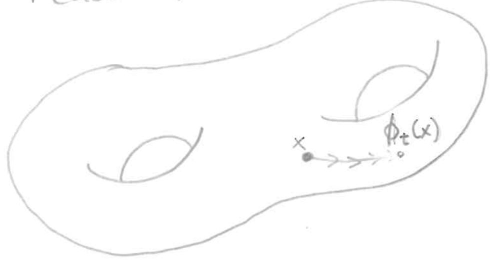


* Dashed line represents an expanding line segment, solid line represents a contracting line segment.

Every $x \in \mathbb{T}^2$ lives on two lines: a line of points that expand and a line of points that contract. (More concretely, look at the behavior of the differential $dA_x : T_x(\mathbb{T}^2) \longrightarrow T_{Ax}(\mathbb{T}^2)$. The differential can be interpreted as $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ acting on $\mathbb{R}^2 \cong T_x(\mathbb{T}^2) \cong T_{Ax}(\mathbb{T}^2)$ linearly, and the eigenspaces correspond to the expanding and contracting directions of this map. Note that these two directions account for a maximal set of linearly independent directions, i.e. there are no directions staying fixed or purely rotating.)

Hopf's observation can be summarized in the following way: for any f that is continuous with some control over the derivatives (eg. Lipschitz), if $f \circ A = f$, the $f(x) = f(y)$ for any x and y lying along the same "contracting line." If A is invertible, the same should be true of expanding lines. Since the collection of expanding lines cover the surface (as does the collection of contracting lines), one should be able to deduce that f is essentially constant. Lastly, being Lipschitz is a dense condition in L^2, so a bootstrapping argument leads us to conclude that $A \circlearrowright (\mathbb{T}^2, dm)$ is ergodic.

NOTE Hopf's argument originally applied to flows. Let $f \in L^2(X, \mu)$ such that $f \circ \phi_t = f$ where ϕ_t is an invertible flow with only expanding or contracting directions. f must be constant in the expanding direction and contracting directions. However, there is good reason to believe that the map $A: \mathbb{T}^2 \to \mathbb{T}^2$, despite being discrete, has the same property for the aforementioned reasons.



We will rigorize this idea soon enough. However, our goal is not merely to show that hyperbolic toral automorphisms are ergodic on $(\mathbb{T}^2, dm)$, we want to observe that they are mixing. To do this, we use an observation that I believe is originally due to Yves Coudène (see "On invariant distributions and mixing," Ergodic Theory and Dynamical Systems, 2006). We will follow Coudène's exposition which you can find in the text "Ergodic Theory and Dynamical Systems" by Coudène. It is probably appropriate to call this style of argument a

"Coudène-Hopf argument." We begin with a few definitions. Note that to make sense of "expanding" and "contracting", it will help to have a metric, so one thing you might want to track is the interplay between the metric space and the underlying (or perhaps "induced" or "corresponding") measure space, which is where we do measurable dynamics.

Let (X,d) be a metric space and $T: X \to X$ an invertible map. Let μ be the Borel measure on (X,d).

<u>DEF</u> Fix $x \in X$. The (strong) <u>stable manifold of $x \in X$ for T</u> is

$$W^{ss}(x) = \{ y \in X : d(T^n(x), T^n(y)) \xrightarrow{n \to \infty} 0 \}.$$

The (strong) <u>unstable manifold of $x \in X$ for T</u> is

$$W^{su}(x) = \{ y \in X : d(T^{-n}(x), T^{-n}(y)) \xrightarrow{n \to \infty} 0 \}.$$

(i.e. the (strong) unstable manifold for T is the (strong) stable manifold for T^{-1}.)

<u>EXERCISE</u> (short!) Show that (strong) stable manifolds partition the space X.

DEF A measurable function $f: X \to \mathbb{R}$ is called <u>W^{ss}-invariant</u> if, after restriction to a full measure set $X_0 \subseteq X$, f is constant on the (strong) stable manifolds:

$$\forall \ x, y \in X_0, \ y \in W^{ss}(x) \Rightarrow f(x) = f(y).$$

Here is the key observation:

THM B Assume $T: X \to X$ preserves μ. Let $f \in L^2(X, \mu)$. The accumulation points (in the weak topology) of the sequence $f \circ T^n$ are W^{ss}-invariant. Furthermore, if T is invertible, then these accumulation points are also W^{su}-invariant.

We will need a bit of machinery to prove this. First, recall that $f \circ T^n \xrightarrow{n \to \infty} g$ in the weak topology if for any $\phi \in L^2(X, \mu)$, $\langle f \circ T^n, \phi \rangle \to \langle g, \phi \rangle$.

(<u>NOTE</u> Some authors call (essentially) this weak-$*$ convergence, e.g. Einsiedler-Ward. Recall that weak-$*$ convergence is a topology on a dual space, eg. $L^n \to L$ in space of linear functionals on V, a normed vector space

the dual of V (called V^*) in the weak-$*$ topology if $L^n(v) \to L(v)$ $\forall v \in V$. If V is reflexive, meaning $(V^*)^* \cong V$, then the weak-$*$ topology on V^{**} coincides with the weak topology on V. Recall that any Hilbert space is reflexive, and that $L^2(X,\mu)$ is a Hilbert space.

Thus, we can think of $L^2(X,\mu)$ with the weak topology, or equivalently, we can think of $L^2(X,\mu)$ as a dual space with the weak-$*$ topology.

<u>EXERCISE</u> If you have never thought through the ideas in the previous paragraph, take a moment to think through them carefully. A helpful reference is Folland's "Real Analysis."

Next, recall the Banach-Alaoglu Theorem: Let V be a normed vector space. Then the closed unit ball in V^* is compact in the weak-$*$ topology.

In our setting ($L^2(X,\mu)$), since the weak-topology and weak-$*$ topology coincide, we can apply Banach-Alaoglu.

Next, we will need functions which have something like derivatives controlled by the distance function on the space, and, with any luck, these functions will be dense in $L^2(X,\mu)$ allowing for a bootstrapping argument. Lipschitz functions are perfect: $f: X \to \mathbb{R}$ is Lipschitz if $\exists\, C$ s.t. for $\forall\, x, y \in X$, $|f(x_2) - f(x_1)| \leq C\, d(x_2, x_1)$.

Lastly, we will need the Banach-Saks Theorem, which I'll recount here:

<u>THM</u> (Banach-Saks)

Let $\mathcal{H}$ be a Hilbert space. Consider elements f and $\{f_n\}_{n\in\mathbb{N}}$ of $\mathcal{H}$ such that $f_n \longrightarrow f$ weakly. Then there exists a subsequence $f_{n_1}, \ldots, f_{n_k}, \ldots$ such that

$$\frac{1}{m} \sum_{k=1}^{m} f_{n_k} \xrightarrow{n\to\infty} f$$

in $\underbrace{\text{norm}}_{\text{also called the strong topology.}}$, i.e. $\|\frac{1}{m} \sum_{k=1}^{m} f_{n_k} - f\| \to 0$.

pf: WLOG, assume $f=0$. We can construct a sub sequence f_{n_k} satisfying $\sum_{i<j} |\langle f_{n_i}, f_{n_j} \rangle| < \infty$. Indeed, we do this by induction. Pick any f_{n_1} and construct f_{n_2} by picking some $f_\ell \in \{f_n\}$ such that $\langle f_{n_1}, f_{n_2} \rangle \le \frac{1}{2^2}$. We can do this since $\langle f_{n_2}, f_n \rangle \xrightarrow{n \to \infty} 0$.

Now, for the inductive step, assume that we have constructed $f_{n_1}, \dots, f_{n_k}$ such that if $i<j$, we have

$$|\langle f_{n_i}, f_{n_j} \rangle| < \frac{1}{2^j}.$$

For all $i \le k$, we have that $\langle f_{n_i}, f_n \rangle \xrightarrow{n \to \infty} 0$, so we can find $f_{n_{k+1}}$ so that

$$|\langle f_{n_i}, f_{n_k} \rangle| \le \frac{1}{2^{k+1}}.$$

That gives us

$$\sum_{i<j} |\langle f_{n_i}, f_{n_k} \rangle| \le \sum_{j=1}^{\infty} \frac{j}{2^j} \le 2.$$

We can now bound $\left\| \frac{1}{m} \sum_{k=1}^{m} f_{n_k} \right\|^2$:

diagonal terms (*) off-diagonal terms

$$\left\| \frac{1}{m} \sum_{k=1}^{m} f_{n_k} \right\|^2 = \underbrace{\frac{1}{m^2} \sum_{k=1}^{m} \| f_{n_k} \|^2} + \frac{2}{m^2} \underbrace{\sum_{1 \le i < j \le m} \langle f_{n_i}, f_{n_j} \rangle}_{\le 2}$$

$$\le \frac{1}{m} \cdot \sup_k \| f_{n_k} \|^2 + \frac{4}{m^2}$$

$$\longrightarrow 0 \quad \text{as } m \to \infty.$$

∎

We will need a few other facts, but we will recall these in the proof of THM B.

pf: (THM B).

Let g be a (weak) accumulation point of $\{f\circ T^n\}_{n\in\mathbb{N}}$.

Then, there exists a subsequence $f\circ T^{n_i} \longrightarrow g$ in the weak topology.

Case 1: Assume f is a bounded Lipschitz function.

Then, there is a further subsequence such that $\frac{1}{m}\sum_{k=1}^{m} f\circ T^{n_{i_k}} \xrightarrow{m\to\infty} g$ in norm.

Moreover, there is a further subsequence $\{\frac{1}{m_\ell}\sum_{k=1}^{m_\ell} f\circ T^{n_{i_k}}\}$ such that

$$\Psi_\ell(x) := \frac{1}{m}\sum_{\ell=1}^{m} f\circ T^{n_{i_{k_\ell}}}(x) \xrightarrow{\ell\to\infty} g(x) \text{ a.e.}$$

(Recall that convergence in norm implies there exist a subsequence converging pointwise almost everywhere!)

So, if $y\in W^{ss}(x)$, then

$$|\Psi_\ell(x) - \Psi_\ell(y)| = \left|\frac{1}{m_\ell}\sum_{k=1}^{m_\ell}\left(f\circ T^{n_{i_k}}(x) - f\circ T^{n_{i_k}}(y)\right)\right|$$

$$\leq \frac{C}{m_\ell} \sum_{k=1}^{m_\ell} d(T^{n_{i_k}}(x), T^{n_{i_k}}(y))$$

$$\longrightarrow 0 \quad \text{as } \ell \to \infty,$$

where in the second to last line, we use that f is Lipschitz, so there exists a $C \geq 0$ s.t. $\forall\, x_1, x_2$,

$$|f(x_2) - f(x_1)| \leq C \cdot d(x_2, x_1),$$

and in the last line, we use that $y \in W^{ss}(x)$ so $d(T^{n_{i_k}}(x), T^{n_{i_k}}(y)) \to 0$. Thus, if $y \in W^{ss}(x)$, since $\Phi_\ell(x) \to g(x)$ a we can conclude that $g(x) = g(y)$. Thus, g is W^{ss}-invariant, as desired.

Case 2: Now let $f \in L^2(X, \mu)$ be arbitrary, and use that Lipschitz functions are dense in L^2. We are again assuming that g is any weak accumulation point of $\{f \circ T^n\}_{n \in \mathbb{N}}$, so there exists a subsequence $\{n_i\}_{i \in \mathbb{N}}$ such that $f \circ T^{n_i} \to g$ weakly.

For any $\varepsilon > 0$, there exists a Lipschitz function $\tilde{f}$ such that $\|f - \tilde{f}\|_2 < \varepsilon$. Now, consider $\{\tilde{f} \circ T^{n_i}\}_{n \in \mathbb{N}}$. Since T is measure-preserving,

$\|\tilde{f}\|_2 = \|\tilde{f} \circ T^{n_i}\|_2$, and we have that $\{\tilde{f} \circ T^{n_i}\}_{n\in\mathbb{N}}$ is a sequence with bounded L^2-norm. Thus, we can apply Banach-Alaoglu (after rescaling, if needed): $\{f \circ T^{n_i}\}_{n\in\mathbb{N}}$ is contained in a compact set in $L^2(X,\mu)$, so there exists a subsequence n_{i_j} such that

$$\tilde{f} \circ T^{n_{i_j}} \xrightarrow{\text{weakly}} \tilde{g}.$$

By case 1, we know that $\tilde{g}$ is W^{ss}-invariant.

Now, consider $(f-\tilde{f}) \circ T^{n_{i_j}}$. We have that

rem: $f \circ T^{n_i} \underset{\text{weakly}}{\longrightarrow} g$.

$$(f-\tilde{f}) \circ T^{n_{i_j}} \xrightarrow[j\to\infty]{\text{weakly}} (g-\tilde{g}).$$

(Use the definition of weak convergence + properties of the inner product.)

Thus,

$(*)$ use Hahn-Banach + weak convergence + def. of op. norm

$$\|g-\tilde{g}\| \overset{(*)}{\leq} \liminf_{j\to\infty} \|(f-\tilde{f}) \circ T^{n_{i_j}}\| = \|f-\tilde{f}\| < \varepsilon.$$

(<u>EXERCISE</u> Rigorously justify $(*)$.)

Consequently, there is a sequence that we can construct $\{\tilde{g}_i\}$ converging to g in norm. By taking a subsequence, we have that

$$\tilde{g}_{i_l} \longrightarrow g \quad \underline{\underline{\text{a.e.}}}$$

$$\|g\|_2^2 = \langle g, g \rangle = \lim_{k\to\infty} \langle \underbrace{f\circ T^{n_{i_k}}}_{``\langle U_{T^{n_{i_k}}}(f),\, g\rangle"}, g \rangle \quad = \langle f, U_{T^{n_{i_k}}}^{-1} g \rangle$$

$$= \lim_{k\to\infty} \langle f, g\circ T^{-n_{i_k}} \rangle$$

$$= \langle f, g_0 \rangle, \quad f \in V^\perp, g_0 \in V$$

$$= 0.$$

Thus, $g = 0$ in L^2, and we conclude that every weak limit of $f\circ T^{n_i}$ is 0, as desired.

To complete Case 3, let $f \in L^2(X,\mu) = \overset{W^{\text{su-inv't.}}}{V} \oplus V^\perp$.

We can write: $f = f_1 + f_2$ for $f_1 \in V$ and $f_2 \in V^\perp$.

By the previous argument $f_2 \circ T^{n_i} \xrightarrow{\text{weakly}} 0$ for any convergent subsequence. Thus, if g is a weak limit of $\{f\circ T^n\}_{n\in\mathbb{N}}$, we have

$$f\circ T^{n_i} = f_1\circ T^{n_i} + \underbrace{f_2\circ T^{n_i}}_{\to 0} \xrightarrow{\text{weakly}} g,$$

i.e. $f_1\circ T^{n_i} \xrightarrow{\text{weakly}} g$. In other words, all weak accumulation points of $\{f \circ T^n\}_{n \in \mathbb{N}}$ are weak accumulation points of $\{f_1 \circ T^n\}_{n \in \mathbb{N}}$. Observe that if $f_1 \in V$, then $f_1 \circ T \in V$, and so these accumulation points must be in V, the subspace of W^{su}-invariant functions.

$\boxtimes$

Now, we have that any weak limit of $f \circ T^n$ must be W^{ss}-invariant, and if T is invertible, must be W^{su}-invariant.

There is a straightforward corollary, which amounts to Hopf's observation:

<u>Cor</u> (Hopf)

Let (X,d) metric space, μ a finite Borel measure, and let $T : X \longrightarrow X$ be a m.p. transformation. Then any T-invariant function $f \in L^2(X,\mu)$ is W^{ss}-invariant. Moreover, if T is invertible, f is W^{su}-invariant.

<u>pf:</u> Observe that f is a weak limit of $\{f \circ T^n\}_{n \in \mathbb{N}}$ ... since $f \circ T^n = f$! Apply <u>THM B</u>.

$\boxtimes$

Thus, since each $\tilde{g}_{ij}$ is W^{ss}-invariant, g must be W^{ss}-invariant. (Else, there would exist a J such that for all $j > J$, $\tilde{g}_{ij}$ could no longer be W^{ss}-invariant!) This completes <u>Case 2</u>.

<u>Case 3</u>: Lastly, we consider when T is invertible. Let $V \subseteq L^2(X, \mu)$ be the subspace of W^{su}-invariant functions. Assume $f \in V^{\perp}$. We will show that $f \circ T^n \longrightarrow 0$ weakly, and consequently, any weak limit of $\tilde{f} \circ T^n$ for $\tilde{f} \in L^2(X, \mu)$ must be contained in V.

Assume g is a weak accumulation point of $\{f \circ T^n\}_{n \in \mathbb{N}}$.
Then there exists a subsequence $f \circ T^{n_i} \xrightarrow{\text{weakly}} g$.

We use T^{-1}: consider $g \circ T^{-n}$. By <u>case 2</u>, there exists a subsequence $g \circ T^{-n_{i_k}} \longrightarrow g_0 \in V$. (Remember, $W^{su}(x)$ is the stable manifold for T^{-1}!)

So, we have

We can apply this argument to hyperbolic toral automorphisms.

<u>THM A</u> Let A be an $n\times n$ matrix with integer coefficients, determinant 1, and no eigenvalues on the unit circle. Then

$$A : \mathbb{T}^n \longrightarrow \mathbb{T}^n \quad , \text{ when } \mathbb{T}^n \cong \mathbb{R}^n/\mathbb{Z}^n$$

preserves the Lebesgue measure and is mixing

pf ② (Hopf Argument): Assume $A : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathbb{R}^n = \hat{E}^s \oplus \hat{E}^u$, where $\hat{E}^s$ is the smallest subspace containing all eigenspaces with eigenvalues $|\lambda| < 1$. Let $\hat{E}^u$ be the smallest vector subspace containing all eigenspaces with eigenvalues $|\lambda| > 1$. Since there are no eigenvalues on the unit circle, $\hat{E}^s \oplus \hat{E}^u = \mathbb{R}^n$. Since $\det(A) = 1$, neither $\hat{E}^s$ or $\hat{E}^u$ is the 0 subspace.

Let E_s and E_u be the projectors of $\hat{E}^s$ and $\hat{E}^u$ to $\mathbb{T}^n \cong \mathbb{R}^n/\mathbb{Z}^n$. Then, the stable and unstable manifolds for a fixed $x \in \mathbb{T}^n$ is given by

$$W^{ss}(x) = x + E_s$$

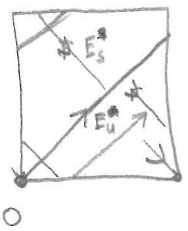
and

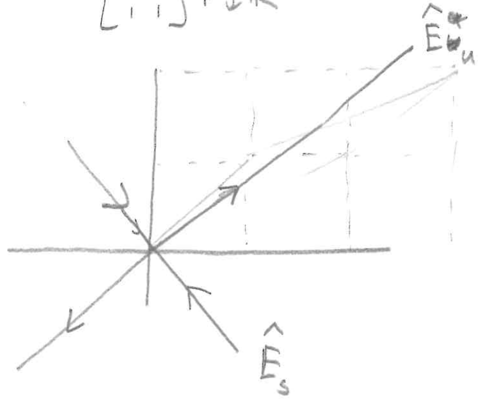
$$W^{su}(x) = x + E_u$$

(Indeed, if $y \in x + E_s$, then $d(A^n y, A^n x) = d(A^n(x+e_s), A^n x)$
$$= d(A^n x + \underbrace{A^n e_s}_{\text{all } \lambda < 0}, A^n x)$$
$$\to 0.$$

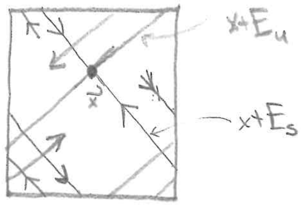
If $y \in W^{ss}(x)$, then $d(A^n y, A^n x) \to 0$. Let $v = y - x$. Then $d(A^n y, A^n x) = d(A^n v + A^n x, A^n x) = d(A^n v, 0) \to 0$, so $v \in E_s$)

• Example $\begin{bmatrix}2&1\\1&1\end{bmatrix} \circlearrowright \mathbb{T}^2$ $\qquad$ $\begin{bmatrix}2&1\\1&1\end{bmatrix} \circlearrowright \mathbb{R}^2$



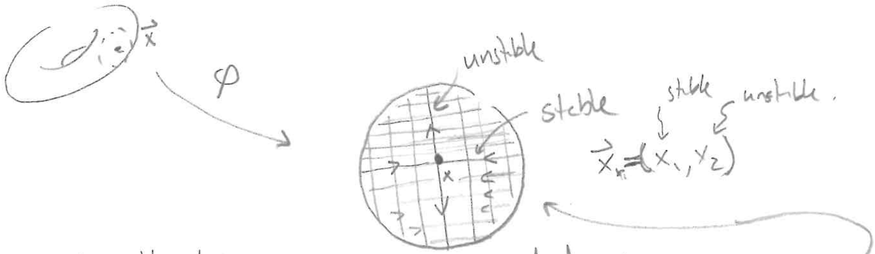


$\circlearrowleft$ at $x \in \mathbb{T}^2$



happen to be right angles .. not necessarily true in general...

• Use E_u and E_s to create coordinate charts around any $\vec{x} \in \mathbb{T}^n$



Moreover, the Lebesgue measure is $dxdy$ in such charts. $\varphi_* m = \det(\varphi)\,dxdy$, pick φ accordingly!

• Let $f \in L^2(\mathbb{T}^n)$ and let g be a weak-limit of $f \circ A^n$.

By the Hopf argument, we know g is W^{ss} and W^{su}-invariant.

(it does not depend on $\vec{x}$ ... in any coordinate chart like the one constructed above, g can be restricted to a full measure subset, and will still be constant along stable/unstable pieces in the chart).

If it is constant in two directions in a full measure set... is it not constant a.e. in these charts? Indeed...

<u>LEMMA</u> Let $(X, \mathcal{B}, \mu)$ and $(Y, \mathcal{C}, \nu)$ be prob. spaces. Let $f: X \times Y \to \mathbb{R}$ be an L^2-function. Assume that there exists two measurable functions $\varphi_1 : X \to \mathbb{R}$ and $\varphi_2 : Y \to \mathbb{R}$

and a subset $Z \subset X \times Y$ of full ~~measure~~ $\mu \otimes \nu$-measure such that

$$\forall (x,y) \in Z, \quad f(x,y) = \varphi_1(x) \quad \text{and} \quad f(x,y) = \varphi_2(y).$$

$\varphi_1(x) = \varphi_2(y)$ in a full measure set...

Then f is constant almost everywhere.

pf: Let $Z_{y_0} = \{x \in X : (x,y) \in Z\}$, $Z_{x_0} = \{y \in Y : (x,y) \in Z\}$.

Full measure.

Then $\mu \otimes \nu(Z) = \int_Y \mu(Z_y)\, d\nu(y) = \int_X \nu(Z_x)\, d\mu(y) \Big] = 1$

Fubini ... really Tonelli ... for characteristics!

$\Rightarrow \exists\, Z_{x_0}$ s.t. $\nu(Z_{x_0}) = 1$ (in fact many!)

So, $x_0 \times Z_{x_0} \subset Z$. Consider any

$$(x,y) \in Z \cap \{X \times Z_{x_0}\}.$$

Know, $(x_0, y) \in Z$, and so

$$\boxed{\varphi(x_0)} = f(x_0, y) = \varphi(y)$$

and

$$\varphi(x) = \boxed{f(x,y)} = \varphi(y)$$

So, $f(x,y) = \varphi(x_0)$. i.e., f is constant a.e.

▨

So, by the Lemma, g is locally (almost) constant. We now need some kind of local-to-global argument to show g is, in fact, constant almost everywhere.

could say separable to remove finite

LEMMA Let (X,d) be a metric space, let μ be a finite measure with connected support. Let g be a locally (almost) constant function. Then g is constant for almost all $x \in \operatorname{supp}(\mu)$.

pf: Since g is locally almost constant, for any $x_0 \in \operatorname{Supp}\mu$, there exists a radius r_{x_0} such that g is constant a.e. on $B_{r_{x_0}}(x_0)$.

Let $g\big|_{B_{r_{x_0}}(x_0)} = C_{x_0}$. Our goal is to show that C_{x_0} does not depend on x_0.

For any $x \in \operatorname{supp}\mu$, define

$$\bar{g}(x) := \overline{\lim_{\substack{r \to 0 \\ (r < r_x)}}} \frac{1}{\mu(B_x(r))} \int_{B(x,r)} g \, d\mu.$$

look inside the Ball where we know g is almost constant? Not needed.. taking a limit by values would have to recur!

not just locally almost constant!

Then $\bar{g}$ is locally constant on $\operatorname{supp}\mu$: indeed, for any $x \in B_{r_{x_0}}(x_0) \cap \operatorname{Supp}\mu$,

$$\bar{g} = \overline{\lim_{\substack{r \to 0 \\ r < r_{x_0}}}} \frac{1}{\mu(B_x(r))} \int_{B(x,r)} C_{x_0} \, d\mu = C_{x_0}.$$

Now, we will use the fact that the support is connected.

[**Claim**: the support of a finite measure has a countable basis.] ← see Coudéne Ch.18 for a proof!

$\exists$ a collection of open sets $\mathcal{D}$ s.t. every non-empty open set can be written as a union of elements $U \in \mathcal{D}$.

For every $x_0 \in \operatorname{supp}\mu$, there is $U_{x_0} \subset B_{r_{x_0}}(x_0)$.

Then $\{U_{x_0}\}_{x_0 \in \operatorname{supp}\mu} \subset \mathcal{D}$. is countable and covers μ.

Thus, $\bar{g}$ is constant on $\operatorname{supp}\mu$. $\xrightarrow{\text{a.e.}}$ $C_{x_0} = C_{y_0}$ for any $x_0, y_0 \in \operatorname{supp}\mu$.

Let $C := C_{x_0}$. Then $g \equiv C_{x_0} = C$ a.e. on $\operatorname{supp}\mu$.

$\blacksquare$

To finish the proof of the theorem, apply this second lemma to deduce that every accumulation point of $f \circ T^n$ is constant almost everywhere.

Notice $$\int_{\mathbb{T}^2} g \, d\mu = \langle g, \chi_{\mathbb{T}^2} \rangle = \lim_{n \to \infty} \langle f \circ A^n \chi_{\mathbb{T}^2} \rangle = \int_{\mathbb{T}^2} f \circ A^n \, d\mu \overset{\text{m.p.t.}}{=} \int_{\mathbb{T}^2} f \, d\mu.$$

Moreover, since g is constant a.e., $\int_{\mathbb{T}^2} g \, d\mu = g \cdot \mu(\mathbb{T}^2) = g$.

Thus, $g = \int_{\mathbb{T}^2} f \, d\mu$. So, any weak limit is equal to $\int_{\mathbb{T}^2} f \, d\mu$:

$$\langle f \circ A^n, \phi \rangle \xrightarrow[n \to \infty]{} \left\langle \int_{\mathbb{T}^2} f \, d\mu, \phi \right\rangle = \int_{\mathbb{T}^2} f \, d\mu \cdot \int_{\mathbb{T}^2} \phi \, d\mu.$$

Thus, A is mixing with respect to the Lebesgue measure. $\boxtimes$