Theorem

Let $A : \mathbb{T}^2 \to \mathbb{T}^2$ by $A(x) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mod 1$.

$A$ is mixing with respect to Lebesgue measure.

Recall that $A$ is mixing if for any pair of measurable set $S_1, S_2$ we have that

$$\lim_{n \to \infty} \text{Leb}(A \cap A^{-n} S_1 \cap S_2) = \text{Leb}(S_1) \text{Leb}(S_2).$$

The proof is much more general.
Two facts

1. If $A$ is not mixing then there exists $n_1, \ldots$ and $f \in L^2$, $g$ non-constant so that

$$\frac{1}{m} \sum_{i=0}^{m} f \circ T^{n_i} \rightarrow g \text{ and } \frac{1}{m} \sum_{i=0}^{m} f \circ T^{-n_i} \rightarrow g$$

almost everywhere.

2. Given $g$ as in fact 1, there exists a set of full measure $G$ so that if $x, y \in G$ and $\lim_{n \to \infty} d(A^n x, A^n y) = 0$ then $g(x) = g(y)$ and if $(x, y) \in G$ and $\lim_{n \to -\infty} d(A^n x, A^n y) = 0$ then $g(x) = g(y)$. 
Proof of Theorem assuming facts

- If \( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + t \begin{pmatrix} 1 \\ \frac{\sqrt{5} - 1}{2} \end{pmatrix} \) \( \mod 1 \) then
  \[ \lim_{n \to -\infty} d(A^n x, A^n y) = 0. \]

- If \( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + t \begin{pmatrix} 1 \\ \frac{\sqrt{5} + 1}{2} \end{pmatrix} \) \( \mod 1 \) then
  \[ \lim_{n \to \infty} d(A^n x, A^n y) = 0. \]

- For almost every \( z \) we have \( z + t \begin{pmatrix} 1 \\ \frac{\sqrt{5} - 1}{2} \end{pmatrix} \) is in \( G \) for a.e. \( t \).

- For almost every \( t \) we have \( z + t \begin{pmatrix} 1 \\ \frac{\sqrt{5} - 1}{2} \end{pmatrix} + s \begin{pmatrix} 1 \\ \frac{\sqrt{5} + 1}{2} \end{pmatrix} \) is in \( G \) for a.e. \( s \).

- For any such \( z \) we have \( g(z') = g(z) \) for a.e. \( z' \).

- So this contradicts fact 1.
**Fact 1**

**Lemma**

If $T : (X, \mu) \to (X, \mu)$ is not mixing then there exist, $f, g \in L^2(\mu)$ with $\int fd\mu = 0$ and $m_k \to \infty$ so that

$$f \circ T^{m_k} \to g \text{ and } f \circ T^{-m_k} \to g$$

weakly in $L^2(\mu)$.

This means that for any $h \in L^2(\mu)$ we have $\langle f \circ T^{m_k}, h \rangle \to \langle g, h \rangle$ and $\langle f \circ T^{-m_k}, h \rangle \to \langle g, h \rangle$. 
Why weak limits? The weak topology on $L^2(\mu)$ of norm at most $N$ is a compact metric space. (If $\{\phi_j\}$ is an orthonormal basis $d(f, g) = \sum_{j=1}^{\infty} |\langle f - g, \phi_j \rangle|$ is such a a metric.)

So we have limits (which may be projection onto constants)!

In fact if $f \in L^2(\mu)$ and $T$ is $\mu$-measure preserving then $f \circ V^i$ along a subsequence. More is true if $T$ is $\mu$ measure preserving then the Koopman operator $U^n_T$ has a subsequential limit in the weak operator topology, the topology of pointwise convergence in the weak topology.

This follows by observing that if $\{\phi_j\}$ is an orthonormal basis and $V_1, \ldots, V_\infty$ are unitary operators so that $\lim_{j \rightarrow \infty} \langle V_j \phi, g \rangle = \langle V_\infty \phi_j, g \rangle$ for all $g$ in $L^2$ then $\langle V_j h, g \rangle = \langle V_\infty h, g \rangle$ for all $g, h$ in $L^2$. 

\[ \lim_{j \rightarrow \infty} \langle V_j \phi, g \rangle = \langle V_\infty \phi_j, g \rangle \]
Weak convergence for mixing transformations

$(X, \mu, T)$ is mixing iff $U^n_T$ converges weakly to projection onto the constant functions, but it does not converge strongly to anything.

Indeed, because the span of characteristic functions of measurable sets is dense in $L^2$, $T$ mixing implies that for all $f, g \in L^2$ we have that

$$\langle f \circ T^n, g \rangle \rightarrow \langle f, 1 \rangle \langle 1, g \rangle.$$

In general, $U^n_T$ does not need to converge to anything in the weak operator topology, but it does converge along a subsequence.
This proof uses the spectral theorem:

**Theorem**

Let $T : L^2(\mu) \to L^2(\mu)$ is unitary and $f \in L^2(\mu)$. Let $H_f = \text{span}\{f \circ T^n\}$. For each $f \in L^2(\mu)$ there exists $\sigma_f$, a measure on $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and $V : H_f \to L^2(\sigma_f)$, unitary, so that

$$
\langle f, f \rangle = \int d\sigma_f \text{ and } \langle g \circ T, f \rangle = \int_{S^1} z \cdot (Vg) d\sigma_f.
$$

- If $f$ is an eigenfunction, with eigenvalue $\lambda$ then $H_f = \mathbb{C}f$ and $\sigma_f$ is point mass at $\lambda$.
- This theorem is convenient for the weak topology, because the statement is about inner products.
- Note that if $g = \sum a_j f \circ T^j$ we have $V(g) = \sum a_j z^j$ (as a function in $L^2(\sigma_f)$).
Proof of Lemma

- There exists such $f$ so that $f \circ T^n \not\to 0$ weakly.
- So there exists $n_j \in \mathbb{Z}$ and $\phi \in L^2(\mu)$ so that $f \circ T^{n_j} \to \psi$ weakly.
- So by the spectral theorem there exists $n_j$ so that $z^{n_j} \to \phi = V\psi$ weakly in $L^2(\sigma_f)$ the spectral measure (on $S^1$) associated to $f$.
- Note that $z^{-n_j} \to \bar{\phi}$ weakly.
- One can show that there exist $m_j, m'_j$ so that $z^{nm_j - nm'_j} \to \phi \cdot \bar{\phi}$ in $L^2(\sigma_f)$. [details]
- Since $\phi \cdot \bar{\phi}$ is real valued, the same is true for $z^{-(nm_j - nm'_j)}$.
- By the spectral theorem if $h = \lim \sum a_j f \circ T^j$ then
  $$\langle f \circ T^n, h \rangle = \lim \sum a_j \int z^{n-j} d\sigma_f.$$  
- So $f \circ T^{nm_j - nm'_j}$ converges weakly on $\overline{\text{span}\{f \circ T^n\}}$.
- It is trivial that it also converges weakly on $\overline{\text{span}\{f \circ T^n\}^\perp}$.  

Theorem
(Banach-Saks) If $f_j \to g$ weakly then there exists $n_j$ so that

$$
\frac{1}{m} \sum_{j=1}^{m} f_{n_j} \to g \text{ a.e.}
$$

- It suffices to show this for $g \equiv 0$.
- Given $n_1, \ldots, n_{j-1}$, choose $n_j$ so that $\langle f_{n_j}, f_{n_\ell} \rangle < \frac{1}{2^j}$ for all $\ell < j$.
- Observe $\langle \sum_{j=1}^{m} f_{n_j}, \sum_{j=1}^{m} f_{n_j} \rangle = O(m)$ and so $\frac{1}{m} \sum_{j=1}^{m} f_{n_j} \to 0$ in $L^2$. [details]
- $L^2$ convergence implies convergence a.e. along a subsequence.
Fact 2

- By Lusin’s theorem, for every $\epsilon > 0$ there exists $k$ so that for all $\ell > k$ there is a set $B_{\ell}$ of measure at most $\epsilon$ so that if $x, y \notin B_{\ell}$ and $d(A^i x, A^i y) \to 0$ then
  \[
  \left| \frac{1}{\ell} \sum_{i=1}^{\ell} f(A^{ni} x) - \frac{1}{\ell} \sum_{i=1}^{\ell} f(A^{ni} y) \right| < \epsilon.
  \]

- This gives that there is a full measure set $G'$ so that if $x, y \in G'$ and $d(A^i x, A^i y) \to 0$ then
  \[
  \left| \frac{1}{\ell} \sum_{i=1}^{\ell} f(A^{ni} x) - \frac{1}{\ell} \sum_{i=1}^{\ell} f(A^{ni} y) \right| \to 0 \text{ along a subsequence.}
  \]

- Because $\frac{1}{m} \sum_{i=1}^{m} f(A^{ni})$ converges almost everywhere we have fact 2.
  -Let $G$ be the full measure set where we have convergence intersected with $G'$. 
Lemma

If $z^{j_\ell} \to g$ and $z^{k_\ell} \to h$ weakly in $L^2(\sigma)$ then there exist subsequence $j'_{\ell}$ and $k'_{\ell}$ so that $z^{j'_{\ell}+k'_{\ell}} \to g \cdot h$ weakly.

- For any $r \in \mathbb{Z}$ and $\phi \in L^2(\sigma)$ we have that
  \[
  \int gz^r \phi d\sigma = \lim_{j_\ell} \langle z^{j_\ell+r}, \phi \rangle.
  \]
- \[
  \langle g \cdot h, \phi \rangle = \langle g, \bar{\phi} \cdot h \rangle = \lim_{\ell \to \infty} \langle z^{j_\ell} \cdot h, \phi \rangle.
  \]
- Now for each $\epsilon > 0$, $\phi$ and $\ell$ there exists $k_0$ so that for all $r \geq r_0$ we have
  \[
  |\langle z^{j_\ell} \cdot z^{k_r}, \phi \rangle - \langle z^{j_\ell} \cdot h, \phi \rangle| < \epsilon.
  \]
- So there exists $n_\ell, m_\ell$ subsequence of $j_\ell$ and $k_\ell$ respectively
  so that $\langle z^{n_\ell} z^{m_\ell}, \phi \rangle \to \langle g \cdot h, \phi \rangle$.
- Choosing $\phi_1, \ldots$ an ortho-normal basis for $L^2(\sigma)$ and applying a
  diagonal argument gives the lemma.
\[ \langle \sum_{j=1}^{m} f_j, \sum_{j=1}^{m} \rangle \leq \sum_{j=1}^{m} \langle f_j, f_j \rangle + \sum_{k=1}^{j-1} 2 \text{Re}(\langle f_j, f_k \rangle) \leq m \sup \{ \| f \|_2^2 \} + 2 \sum_{j=1}^{m} \frac{j-1}{2j} \]