MATH- 62052/72052
Functions of Real Variables 2.
Lecture 25.

Artem Zvavitch

Department of Mathematical Sciences, Kent State University

Spring, 2020

We will play in σ-finite measure space $(X, \mathcal{M}, \mu)$. One of our heroes will be

Measure-preserving transformation:

A mapping
$$\tau : X \to X \text{ such that } \mu(\tau^{-1}E) = \mu(E) \text{ for all } E \in \mathcal{M},$$
here $\tau^{-1}(E)$ is a pre-image of E, i.e. $\tau^{-1}(E) = \{x \in X : \tau(x) \in E\}$.

Also if τ is measure-preserving transformation + bijection + τ^{-1} is measure-preserving transformation, then τ is **measure-preserving isomorphism**.

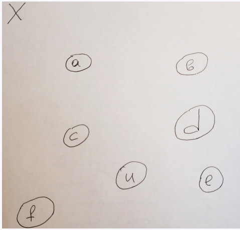
Very basic fact

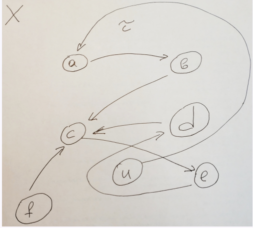
τ - measure preserving, f- measurable, then $f(\tau(x))$ is measurable and
$$\int_X f(\tau(x))d\mu(x) = \int_X f(x)d\mu(x).$$

$$\mu(\{x \in X : f(\tau(x)) > t\}) = \mu(\{y \in X : f(y) > t\}).$$

Transformations and time

We can imagine our space X to be a set of "states" of some system S.

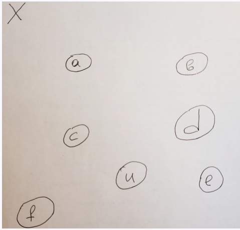


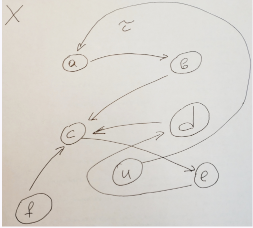


At each (discrete) time n our system at some state $x \in X$ (i.e. $S(n) = x$). What we know that at time $n+1$ our system is at state $S(n+1) = \tau(S(n))$, where the mapping $\tau : X \to X$ presents a transformation of our system after a unit of time, i.e. the next step of the system $(S(n+1))$ depends only on "position" of the system $(S(n))$ and independent of the moment of time when the step is taken.

Transformations and time

We can imagine our space X to be a set of "states" of some system S.





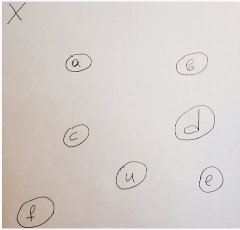
At each (discrete) time n our system at some state $x \in X$ (i.e. $S(n) = x$). What we know that at time $n+1$ our system is at state $S(n+1) = \tau(S(n))$, where the mapping $\tau : X \to X$ presents a transformation of our system after a unit of time, i.e. the next step of the system $(S(n+1))$ depends only on "position" of the system $(S(n))$ and independent of the moment of time when the step is taken.
So if we started at point $x \in S$ then after k steps we will be at point

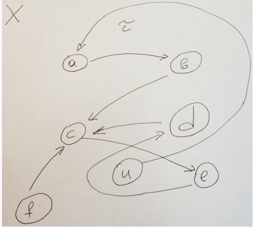
$$\tau \circ \tau \circ \cdots \circ \tau(x) = \tau^k(x).$$

Thus the mapping τ^k describe the evolution of the system after k units of time. It is interesting to understand the limiting behavior of the system as well as the average behavior of the system and different parameters associated to it.

Transformations and time

We can imagine our space X to be a set of "states" of some system S.





At each (discrete) time n our system at some state $x \in X$ (i.e. $S(n) = x$). What we know that at time $n+1$ our system is at state $S(n+1) = \tau(S(n))$, where the mapping $\tau : X \to X$ presents a transformation of our system after a unit of time, i.e. the next step of the system ($S(n+1)$) depends only on "position" of the system ($S(n)$) and independent of the moment of time when the step is taken.
So if we started at point $x \in S$ then after k steps we will be at point

$$\tau \circ \tau \circ \cdots \circ \tau(x) = \tau^k(x).$$

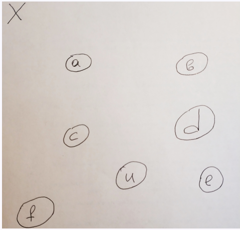
Thus the mapping τ^k describe the evolution of the system after k units of time. It is interesting to understand the limiting behavior of the system as well as the average behavior of the system and different parameters associated to it. Last time we discussed

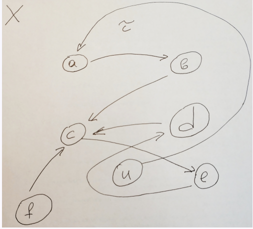
$$A_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)),$$

as well as the limits of $A_n(f)(x)$ as $n \to \infty$ and applied what we learned about Hilbert spaces.

Transformations and time

We can imagine our space X to be a set of "states" of some system S.





At each (discrete) time n our system at some state $x \in X$ (i.e. $S(n) = x$). What we know that at time $n+1$ our system is at state $S(n+1) = \tau(S(n))$, where the mapping $\tau : X \to X$ presents a transformation of our system after a unit of time, i.e. the next step of the system $(S(n+1))$ depends only on "position" of the system $(S(n))$ and independent of the moment of time when the step is taken.
So if we started at point $x \in S$ then after k steps we will be at point

$$\tau \circ \tau \circ \cdots \circ \tau(x) = \tau^k(x).$$

Thus the mapping τ^k describe the evolution of the system after k units of time. It is interesting to understand the limiting behavior of the system as well as the average behavior of the system and different parameters associated to it. Today we will apply "maximal function techniques" and study

$$f^*(x) = \sup_{1 \le m < \infty} A_m(f) = \sup_{1 \le m < \infty} \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x)),$$

Two theorems we proved

Consider a Hilber space $\mathcal{H} = L^2(X, \mathcal{M}, \mu)$. Let $\tau : X \to X$ be a measure preserving transformation. Define a linear operator $T(f)(x) = f(\tau(x))$.
Reminder T is isometry, indeed $\|Tf\|^2 = \int_x f^2(\tau(x))d\mu(x) = \int_x f^2(x)d\mu(x) = \|f\|^2$. Rewrite:

$$A_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) = \frac{1}{n}\sum_{k=0}^{n-1} T^k f(x).$$

Two theorems we proved

Consider a Hilber space $\mathcal{H} = L^2(X, \mathcal{M}, \mu)$. Let $\tau : X \to X$ be a measure preserving transformation. Define a linear operator $T(f)(x) = f(\tau(x))$.
Reminder T is isometry, indeed $\|Tf\|^2 = \int_x f^2(\tau(x))d\mu(x) = \int_x f^2(x)d\mu(x) = \|f\|^2$. Rewrite:

$$A_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) = \frac{1}{n}\sum_{k=0}^{n-1} T^k f(x).$$

Mean Ergodic Theorem

Suppose T is an isometry of the Hilbert space $\mathcal{H}$, and let P be the orthogonal projection on the subspace of the invariant vectors of T. Let

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

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

Two theorems we proved

Consider a Hilber space $\mathcal{H} = L^2(X, \mathcal{M}, \mu)$. Let $\tau : X \to X$ be a measure preserving transformation. Define a linear operator $T(f)(x) = f(\tau(x))$.
Reminder T is isometry, indeed $\|Tf\|^2 = \int_x f^2(\tau(x))d\mu(x) = \int_x f^2(x)d\mu(x) = \|f\|^2$. Rewrite:

$$A_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) = \frac{1}{n}\sum_{k=0}^{n-1} T^k f(x).$$

Mean Ergodic Theorem

Suppose T is an isometry of the Hilbert space $\mathcal{H}$, and let P be the orthogonal projection on the subspace of the invariant vectors of T. Let

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

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

Maximal function approach: $f^*(x) = \sup\limits_{1 \le m < \infty} A_m(f) = \sup\limits_{1 \le m < \infty} \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$.

Two theorems we proved

Consider a Hilber space $\mathcal{H} = L^2(X, \mathcal{M}, \mu)$. Let $\tau : X \to X$ be a measure preserving transformation. Define a linear operator $T(f)(x) = f(\tau(x))$.
Reminder T is isometry, indeed $\|Tf\|^2 = \int_x f^2(\tau(x))d\mu(x) = \int_x f^2(x)d\mu(x) = \|f\|^2$. Rewrite:

$$A_n(f)(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k(x)) = \frac{1}{n}\sum_{k=0}^{n-1} T^k f(x).$$

Mean Ergodic Theorem

Suppose T is an isometry of the Hilbert space $\mathcal{H}$, and let P be the orthogonal projection on the subspace of the invariant vectors of T. Let

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

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

Maximal function approach: $f^*(x) = \sup\limits_{1 \le m < \infty} A_m(f) = \sup\limits_{1 \le m < \infty} \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$.

Assume $f \in L^1(X, \mu)$ and τ is a measure preserving transformation, then the maximal function $f^*(x)$ is finite for almost every x. Moreover,

$$\mu(\{x : f^*(x) > \alpha\}) \le \frac{6}{\alpha}\|f\|_{L^1(X,\mu)}, \text{ for all } \alpha > 0.$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Remark: In light of the previous slide we notice that if $f \in L^2(X, \mu)$ then the limit must be $P(f)$, where P be the orthogonal projection on the subspace of the invariant vectors of operator $Tf = f(\tau(x))$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Remark: In light of the previous slide we notice that if $f \in L^2(X, \mu)$ then the limit must be $P(f)$, where P be the orthogonal projection on the subspace of the invariant vectors of operator $Tf = f(\tau(x))$. In more general settings if we denote the limit by $\tilde{P}(f)$ we will get

$$\int_X |\tilde{P}(f)(x)| d\mu(x) \leq \int_X |f(x) d\mu(x).$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Remark: In light of the previous slide we notice that if $f \in L^2(X, \mu)$ then the limit must be $P(f)$, where P be the orthogonal projection on the subspace of the invariant vectors of operator $Tf = f(\tau(x))$. In more general settings if we denote the limit by $\tilde{P}(f)$ we will get

$$\int_X |\tilde{P}(f)(x)| d\mu(x) \le \int_X |f(x) d\mu(x).$$

Indeed, if $f \in L^1(X)$ then

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

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Remark: In light of the previous slide we notice that if $f \in L^2(X, \mu)$ then the limit must be $P(f)$, where P be the orthogonal projection on the subspace of the invariant vectors of operator $Tf = f(\tau(x))$. In more general settings if we denote the limit by $\tilde{P}(f)$ we will get

$$\int_X |\tilde{P}(f)(x)| d\mu(x) \le \int_X |f(x) d\mu(x).$$

Indeed, if $f \in L^1(X)$ then

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

But (from the Theorem) $A_m(f)(x)$ converges to $\tilde{P}(f)$ almost everywhere and we can apply the Fatou's lemma.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Idea of the Proof: We note that because $\mu(X) = 1$ we get that $L^2(X, \mu) \subset L^1(X, \mu)$,

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Idea of the Proof: We note that because $\mu(X) = 1$ we get that $L^2(X, \mu) \subset L^1(X, \mu)$, indeed

$$\|f\|_{L^1(X,\mu)} = \int_X |f(x)| d\mu(x) \leq \sqrt{\int_X |f(x)|^2 d\mu(x)} \sqrt{\int_X 1^2 d\mu(x)} = \|f\|_{L^2(X,\mu)}.$$

We will consider probability space $(X,\mathcal{M},\mu)$ (i.e. $\mu(X)=1$).

Pointwise Ergodic Theorem

Consider $f\in L^1(X,\mu)$. Then for almost evert $x\in X$ we have that $A_m(f)(x)=\frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m\to\infty$

Idea of the Proof: We note that because $\mu(X)=1$ we get that $L^2(X,\mu)\subset L^1(X,\mu)$, indeed

$$\|f\|_{L^1(X,\mu)}=\int_X |f(x)|d\mu(x)\le\sqrt{\int_X |f(x)|^2 d\mu(x)}\sqrt{\int_X 1^2 d\mu(x)}=\|f\|_{L^2(X,\mu)}.$$

Moreover, $L^2(X,\mu)$ is dense in $L^1(X,\mu)$ (i.e. "near" every $f\in L^1(X,\mu)$ we can find a function from $L^2(X,\mu)$).

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Idea of the Proof: We note that because $\mu(X) = 1$ we get that $L^2(X, \mu) \subset L^1(X, \mu)$, indeed

$$\|f\|_{L^1(X,\mu)} = \int_X |f(x)| d\mu(x) \leq \sqrt{\int_X |f(x)|^2 d\mu(x)} \sqrt{\int_X 1^2 d\mu(x)} = \|f\|_{L^2(X,\mu)}.$$

Moreover, $L^2(X, \mu)$ is dense in $L^1(X, \mu)$ (i.e. "near" every $f \in L^1(X, \mu)$ we can find a function from $L^2(X, \mu)$). Indeed if $f \in L^1(X, \mu)$ consider a sequence

$$f_n(x) = \begin{cases} f(x) & |f(x)| \leq n \\ 0 & |f(x)| \geq n. \end{cases}$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Idea of the Proof: We note that because $\mu(X) = 1$ we get that $L^2(X, \mu) \subset L^1(X, \mu)$, indeed

$$\|f\|_{L^1(X,\mu)} = \int_X |f(x)| d\mu(x) \leq \sqrt{\int_X |f(x)|^2 d\mu(x)}\sqrt{\int_X 1^2 d\mu(x)} = \|f\|_{L^2(X,\mu)}.$$

Moreover, $L^2(X, \mu)$ is dense in $L^1(X, \mu)$ (i.e. "near" every $f \in L^1(X, \mu)$ we can find a function from $L^2(X, \mu)$). Indeed if $f \in L^1(X, \mu)$ consider a sequence

$$f_n(x) = \begin{cases} f(x) & |f(x)| \leq n \\ 0 & |f(x)| \geq n. \end{cases}$$

Then $\int_X |f(x)|^2 d\mu(x) \leq \int_X n^2 d\mu(x) \leq n^2$ and this $f_n \in L^2(X, \mu)$ and $\|f - f_n\|_1 \to 0$ by dominated convergence theorem (indeed $f(x) - f_n(x) \to 0$ almost every where and $|f(x) - f_n(x)| \leq f(x)$.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Idea of the Proof: We note that because $\mu(X) = 1$ we get that $L^2(X, \mu) \subset L^1(X, \mu)$, indeed

$$\|f\|_{L^1(X,\mu)} = \int_X |f(x)| d\mu(x) \le \sqrt{\int_X |f(x)|^2 d\mu(x)}\sqrt{\int_X 1^2 d\mu(x)} = \|f\|_{L^2(X,\mu)}.$$

Moreover, $L^2(X, \mu)$ is dense in $L^1(X, \mu)$ (i.e. "near" every $f \in L^1(X, \mu)$ we can find a function from $L^2(X, \mu)$). Indeed if $f \in L^1(X, \mu)$ consider a sequence

$$f_n(x) = \begin{cases} f(x) & |f(x)| \le n \\ 0 & |f(x)| \ge n. \end{cases}$$

Then $\int_X |f(x)|^2 d\mu(x) \le \int_X n^2 d\mu(x) \le n^2$ and this $f_n \in L^2(X, \mu)$ and $\|f - f_n\|_1 \to 0$ by dominated convergence theorem (indeed $f(x) - f_n(x) \to 0$ almost every where and $|f(x) - f_n(x)| \le f(x)$. So we will prove the theorem for $f \in L_2(X, \mu)$ and after use maximal theorem to push the result to actually true for all functions in $L_1(X, \mu)$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$).

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$). Let us prove that this decomposition is possible. Indeed, we can write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \varepsilon/2$.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$). Let us prove that this decomposition is possible. Indeed, we can write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \varepsilon/2$. To decompose f_1 we need to bring back the geometric lemma we had about properties of T.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$). Let us prove that this decomposition is possible. Indeed, we can write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \varepsilon/2$. To decompose f_1 we need to bring back the geometric lemma we had about properties of T. We proved that if $S \subset L^2$ is a subspace of invariant vectors of T and $S_1 \subset L^2$ is a subspace of the vectors of the form $(I - T)G$, for some $G \in L^2$ then $L^2 = S \oplus \bar{S}_1$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$). Let us prove that this decomposition is possible. Indeed, we can write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \varepsilon/2$. To decompose f_1 we need to bring back the geometric lemma we had about properties of T. We proved that if $S \subset L^2$ is a subspace of invariant vectors of T and $S_1 \subset L^2$ is a subspace of the vectors of the form $(I - T)G$, for some $G \in L^2$ then $L^2 = S \oplus \bar{S}_1$. Thus we can write

$$f' = F_o + F_1 + h,$$

where $F_0 \in S$, $F_1 \in S_1$ (so $F_1 = (I - T)G$) and h is a "payment" for takins F_1 from S_1 and not its closure and thus $\|h\|_{L^2} \leq \varepsilon/2$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$ (remember T is our old friend - $T(f)(x) = f(\tau(x))$). Let us prove that this decomposition is possible. Indeed, we can write $f = f' + h'$, where $f' \in L^2$ and $\|h'\|_{L^1} < \varepsilon/2$. To decompose f_1 we need to bring back the geometric lemma we had about properties of T. We proved that if $S \subset L^2$ is a subspace of invariant vectors of T and $S_1 \subset L^2$ is a subspace of the vectors of the form $(I - T)G$, for some $G \in L^2$ then $L^2 = S \oplus \bar{S}_1$. Thus we can write

$$f' = F_o + F_1 + h,$$

where $F_0 \in S$, $F_1 \in S_1$ (so $F_1 = (I - T)G$) and h is a "payment" for takins F_1 from S_1 and not its closure and thus $\|h\|_{L^2} \le \varepsilon/2$ and thus (do not forget that $\mu(X) = 1$ so $\|h\|_{L^1} \le \|h\|_{L^2} \le \varepsilon/2$) we are done:

$$f = f' + h' = F_o + F_1 + (h + h') = F_o + (I - T)G + H,$$

where $H = h + h'$ and $\|H\|_{L^1} \le \|h\|_{L^1} + \|h'\|_{L^1} \le \varepsilon$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right).$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right).$$

Next, we notice that $\frac{1}{m}T^m G(x)$ converges to zero as $m \to \infty$ for almost all x.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right).$$

Next, we notice that $\frac{1}{m}T^m G(x)$ converges to zero as $m \to \infty$ for almost all x. Indeed, by monotone converges theorem the series $\sum_{m=1}^{\infty} \frac{1}{m^2}(T^m G(x))^2$ converges for each x (may be to ∞),

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right).$$

Next, we notice that $\frac{1}{m}T^m G(x)$ converges to zero as $m \to \infty$ for almost all x. Indeed, by monotone converges theorem the series $\sum_{m=1}^{\infty}\frac{1}{m^2}(T^m G(x))^2$ converges for each x (may be to ∞), but its integral is

$$\int_X \sum_{m=1}^{\infty}\frac{1}{m^2}(T^m G(x))^2 d\mu = \sum_{m=1}^{\infty}\frac{1}{m^2}\int_X G^2(\tau^m(x))d\mu = \sum_{m=1}^{\infty}\frac{1}{m^2}\int_X G^2(x)d\mu = \|G\|^2\sum_{m=1}^{\infty}\frac{1}{m^2} < \infty.$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right).$$

Next, we notice that $\frac{1}{m}T^m G(x)$ converges to zero as $m \to \infty$ for almost all x. Indeed, by monotone converges theorem the series $\sum_{m=1}^{\infty}\frac{1}{m^2}(T^m G(x))^2$ converges for each x (may be to ∞), but its integral is

$$\int_X \sum_{m=1}^{\infty}\frac{1}{m^2}(T^m G(x))^2 d\mu = \sum_{m=1}^{\infty}\frac{1}{m^2}\int_X G^2(\tau^m(x))d\mu = \sum_{m=1}^{\infty}\frac{1}{m^2}\int_X G^2(x)d\mu = \|G\|^2\sum_{m=1}^{\infty}\frac{1}{m^2} < \infty.$$

Thus $\sum_{m=1}^{\infty}\frac{1}{m^2}(T^m G(x))^2$ is finite almost everywhere and $\frac{1}{m}T^m G(x)$ converges to zero almost everywhere.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m} \sum_{k=0}^{m-1} (I - T)T^k G = F_0 + \frac{1}{m} \left((I - T^m)G \right)$$

Next we notice that $\frac{1}{m} T^m G(x)$ converges to zero as $m \to \infty$ for almost all x.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. Now lets apply A_m:

$$A_m(F) = A_m(F_0) + A_m((I - T)G) = F_0 + \frac{1}{m}\sum_{k=0}^{m-1}(I - T)T^k G = F_0 + \frac{1}{m}\left((I - T^m)G\right)$$

Next we notice that $\frac{1}{m}T^m G(x)$ converges to zero as $m \to \infty$ for almost all x. It is also easy to see that $\frac{1}{m}G(x)$ converges to zero for a.e. x and thus $A_m(F)$ converges to F_0 for a.e. x.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$ (after this we can use this fact for $\alpha = 1/k$ take a union over $k = 1, 2, \dots$ still get measure zero and use Cauchy criterion for $A_n(f)(x)$ on the remaining set).

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$ (after this we can use this fact for $\alpha = 1/k$ take a union over $k = 1, 2, \dots$ still get measure zero and use Cauchy criterion for $A_n(f)(x)$ on the remaining set). We can write

$$\begin{aligned}\alpha < &|A_n(f)(x) - A_m(f)(x)| = |(A_n(F)(x) - A_m(F)(x)) + (A_n(H)(x) - A_m(H)(x))| \\ \leq &|(A_n(F)(x) - A_m(F)(x))| + |(A_n(H)(x) - A_m(H)(x))|.\end{aligned}$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$ (after this we can use this fact for $\alpha = 1/k$ take a union over $k = 1, 2, \dots$ still get measure zero and use Cauchy criterion for $A_n(f)(x)$ on the remaining set). We can write

$$\begin{aligned}\alpha <& |A_n(f)(x) - A_m(f)(x)| = |(A_n(F)(x) - A_m(F)(x)) + (A_n(H)(x) - A_m(H)(x))| \\ \leq& |(A_n(F)(x) - A_m(F)(x))| + |(A_n(H)(x) - A_m(H)(x))|.\end{aligned}$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m} \sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$ (after this we can use this fact for $\alpha = 1/k$ take a union over $k = 1, 2, \dots$ still get measure zero and use Cauchy criterion for $A_n(f)(x)$ on the remaining set). We can write

$$\begin{aligned}\alpha < &|A_n(f)(x) - A_m(f)(x)| = |(A_n(F)(x) - A_m(F)(x)) + (A_n(H)(x) - A_m(H)(x))| \\ \leq &|(A_n(F)(x) - A_m(F)(x))| + |(A_n(H)(x) - A_m(H)(x))|.\end{aligned}$$

We know that $A_n(F)$ is already convergent almost everywhere, thus satisfy Cauchy criterium almost everywhere, so E_α is contained (up to measure zero) in

$$E'_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}$$

and to finish the proof we need to estimate that $\mu(E'_\alpha)$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$. Let

$$E'_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}$$

then $\mu(E_\alpha) \leq \mu(E'_\alpha)$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$. Let

$$E'_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}$$

then $\mu(E_\alpha) \leq \mu(E'_\alpha)$. We can apply maximal function estimate to get a bound on $\mu(E'_\alpha)$.

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N\to\infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$. Let

$$E'_\alpha = \{x : \lim_{N\to\infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}$$

then $\mu(E_\alpha) \leq \mu(E'_\alpha)$. We can apply maximal function estimate to get a bound on $\mu(E'_\alpha)$. Indeed

$$\mu(\{x : \lim_{N\to\infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}) \leq \mu(\{x : 2\sup |A_m(H)(x)| > \alpha\})$$

$$= \mu(\{x : \sup |A_m(H)(x)| > \alpha/2\}) \leq \frac{6}{\alpha/2} \|H\|_{L_1} \leq \frac{12\varepsilon}{\alpha},$$

Now we are ready for pointwise convergence.

We will consider probability space $(X, \mathcal{M}, \mu)$ (i.e. $\mu(X) = 1$).

Pointwise Ergodic Theorem

Consider $f \in L^1(X, \mu)$. Then for almost evert $x \in X$ we have that $A_m(f)(x) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$

Proof: We have seen that every function $f \in L^1(X, \mu)$ can be approximated by functions from $L^2(X, \mu)$. We need to be a bit more precise and constructive. We claim that for every $\varepsilon > 0$ we can write

$$f = F + H, \text{ where } \|H\|_{L^1} < \varepsilon, \text{ and } F = F_0 + (I - T)G,$$

where $F_0, G \in L^2$, $T(F_0) = F_0$. (remember T is our old friend - $T(f)(x) = f(\tau(x))$. $A_m(F)$ converges to F_0 for a.e. x. To prove convergence for $A_m f(x)$ define

$$E_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(f)(x) - A_m(f)(x)| > \alpha\}.$$

Our goal is to show that $\mu(E_\alpha) = 0$ for all $\alpha > 0$. Let

$$E'_\alpha = \{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}$$

then $\mu(E_\alpha) \leq \mu(E'_\alpha)$. We can apply maximal function estimate to get a bound on $\mu(E'_\alpha)$. Indeed

$$\mu(\{x : \lim_{N \to \infty} \sup_{n,m \geq N} |A_n(H)(x) - A_m(H)(x)| > \alpha\}) \leq \mu(\{x : 2\sup |A_m(H)(x)| > \alpha\})$$

$$= \mu(\{x : \sup |A_m(H)(x)| > \alpha/2\}) \leq \frac{6}{\alpha/2}\|H\|_{L_1} \leq \frac{12\varepsilon}{\alpha},$$

where ε is arbitrary so $\mu(E_\alpha) = 0$.