

Lecture 24

Birkhoff's ergodic theorem

Graduate Real Analysis

Fall Semester

Measure-preserving systems

- ① **The integer shift system:** $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), |\cdot|, S)$ with $S : \mathbb{Z} \rightarrow \mathbb{Z}$ given by

$$S(x) = x + 1.$$

- ② **The circle rotation system:** $(\mathbb{T}, \mathcal{L}(\mathbb{T}), dx, R_\alpha)$ with the rotation map $R_\alpha : \mathbb{T} \rightarrow \mathbb{T}$ by $R_\alpha(x) = x + \alpha \pmod{1}$ for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

- ③ **The circle-doubling system:** $(\mathbb{T}, \mathcal{L}(\mathbb{T}), dx, D_2)$ with the doubling map $D_2 : \mathbb{T} \rightarrow \mathbb{T}$ given by $D_2(x) = 2x \pmod{1}$.

- ④ **The continued fraction system:** $([0, 1), \mathcal{L}([0, 1)), \mu, T)$ with the Gauss measure

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x},$$

and continued fraction map $T : [0, 1) \rightarrow [0, 1)$ given by $T(0) = 0$ and

$$T(x) = \frac{1}{x} \pmod{1}, \quad \text{when } x \neq 0.$$

Pointwise convergence

Proposition

Let $(T_k)_{k \in \mathbb{N}}$ be a sequence of linear operators $T_k : L^p(X) \rightarrow L^p(X)$ on a σ -finite measure space $(X, B(X), \mu)$. If

$$C(\alpha) = \sup_{f \in L^p(X)} \mu(\{x \in X : \sup_{k \in \mathbb{N}} |T_k f(x)| > \alpha \|f\|_{L^p}\}) \xrightarrow{\alpha \rightarrow \infty} 0,$$

then

$$\mathcal{L}_X^p = \{f \in L^p(X) : \lim_{n \rightarrow \infty} T_n f(x) \text{ exists } \mu\text{-almost everywhere on } X\}$$

is closed in $L^p(X)$.

Proof. Define

$$\Omega(f)(x) = \limsup_{m, n \rightarrow \infty} |T_n f(x) - T_m f(x)|.$$

Proof

We have to show that $\mu(\{x \in X : \Omega(f)(x) > 0\}) = 0$ for all $f \in \overline{\mathcal{L}_X^p}^{\|\cdot\|_{L^p}}$.

It is clear that

$$\Omega(f)(x) \leq 2T_*f(x).$$

Thus,

$$\mu(\{x \in X : \Omega f(x) > \alpha \|f\|_{L^p(X)}\}) \leq C(\alpha/2).$$

For every $\phi \in \mathcal{L}_X^p$, we have that $\lim_{n \rightarrow \infty} T_n \phi(x)$ exists μ -a.e. on X . Thus, $\Omega(\phi)(x) = 0$ μ -a.e. on X and $\Omega(f - \phi)(x) = \Omega(f)(x)$. It follows that

$$\mu(\{x \in X : \Omega(f)(x) > \alpha \|f - \phi\|_{L^p(X)}\}) \leq C(\alpha/2).$$

Now let $f \in \overline{\mathcal{L}_X^p}^{\|\cdot\|_{L^p}}$ and, for a given $\varepsilon > 0$, pick $\phi \in \mathcal{L}_X^p$ so that $\|f - \phi\|_{L^p} \leq \varepsilon^2$. Take $\alpha = \varepsilon^{-1}$ and note that

$$\mu(\{x \in X : \Omega(f)(x) > \varepsilon\}) \leq C((2\varepsilon)^{-1}).$$

Therefore, $\mu(\{x \in X : \Omega(f)(x) > 0\}) = 0$.



Strategy of proving pointwise convergence

In view of the previous proposition, pointwise convergence problems for a sequence of operators $(T_n)_{n \in \mathbb{N}}$ are reduced to a **two-step procedure**:

- 1 In the **first step**, we have to show that the **maximal function**

$$T_* f(x) = \sup_{k \in \mathbb{N}} |T_k f(x)| \text{ satisfies}$$

$$C(\alpha) = \sup_{f \in L^p(X)} \mu(\{x \in X : \sup_{k \in \mathbb{N}} |T_k f(x)| > \alpha \|f\|_{L^p}\}) \xrightarrow{\alpha \rightarrow \infty} 0.$$

In view of the previous proposition, this ensures that the set of $L^p(X)$ functions for which we have pointwise convergence is closed in $L^p(X)$.

- 2 In the **second step**, we have to find a **dense class** $\mathcal{D}$ of functions in $L^p(X)$ for which we have pointwise convergence. In other words, if

$$\overline{\mathcal{D}}^{\|\cdot\|_{L^p}} = L^p(X) \text{ and } \mathcal{D} \subseteq \mathcal{L}_X^p = \overline{\mathcal{L}_X^p}^{\|\cdot\|_{L^p}}, \text{ then } \mathcal{L}_X^p = L^p(X).$$

Combining these two steps we conclude that the pointwise convergence for $(T_n f)_{n \in \mathbb{N}}$ holds for all $f \in L^p(X)$, since

$$\mathcal{L}_X^p = L^p(X).$$

Baire's category theorem

Theorem (Baire's category theorem)

Assume that (X, ρ) is a complete metric space. If $(U_n)_{n \in \mathbb{N}}$ is a countable family of open dense sets in X , then $\bigcap_{n \in \mathbb{N}} U_n$ is dense in X .

Proof. Let V be a nonempty open set in X . We show that

$$V \cap \bigcap_{n \in \mathbb{N}} U_n \neq \emptyset$$

- Since U_1 is dense, $U_1 \cap V \neq \emptyset$. Choose an open ball B_1 of diameter < 1 such that $\text{cl}(B_1) \subseteq U_1 \cap V$.
- Since U_1 is open and dense, by the same argument we get an open ball B_2 of diameter $< 1/2$ such that $\text{cl}(B_2) \subseteq U_1 \cap B_1$.
- Proceeding similarly, we define a sequence $(B_n)_{n \in \mathbb{N}}$ of open balls in X such that for each $n \in \mathbb{N}$, we have
 - (a) $\text{diam}(B_n) < 2^{-n-1}$;
 - (b) $\text{cl}(B_1) \subseteq U_1 \cap V$;
 - (c) $\text{cl}(B_{n+1}) \subseteq U_n \cap B_n$.

Proof

Since (X, ρ) is a complete metric space, by Cantor's theorem

$$\bigcap_{n \in \mathbb{N}} B_n = \bigcap_{n \in \mathbb{N}} \text{cl}(B_n) = \{x\} \quad \text{for some } x \in X.$$

Clearly,

$$x \in V \cap \bigcap_{n \in \mathbb{N}} U_n$$

showing that $V \cap \bigcap_{n \in \mathbb{N}} U_n \neq \emptyset$ as desired. □

Remark

- Baire's theorem is a very important result in mathematics, which is often used in analysis to prove results that have existential nature.
- The name of this theorem comes from Baire's terminology for sets.
- If X is a topological space, a set $E \subseteq X$ is of the **first category**, according to Baire, if E is a **countable union of nowhere dense sets**; otherwise E is of the **second category**.

Dual form of Baire's category theorem

Corollary

Every completely metrizable space is of the second category in itself.

Proof. Let X be a completely metrizable space. Suppose X is of the first category in itself. Choose a sequence $(F_n)_{n \in \mathbb{N}}$ of closed and nowhere dense sets such that $X = \bigcup_{n \in \mathbb{N}} F_n$. Then the sets $U_n = F_n^c$ are dense and open, and $\bigcap_{n \in \mathbb{N}} U_n = \emptyset$. This contradicts the Baire category theorem. $\square$

Theorem (Dual form of Baire's category theorem)

If $(F_n)_{n \in \mathbb{N}}$ is a countable family of nowhere dense sets in a complete metric space (X, ρ) , then $\bigcup_{n \in \mathbb{N}} F_n$ has empty interior.

Proof. Since $\text{int}(\text{cl}(F_n)) = \emptyset$ for every $n \in \mathbb{N}$ observe

$$\text{int}\left(\bigcup_{n \in \mathbb{N}} F_n\right) \subseteq \text{int}\left(\bigcup_{n \in \mathbb{N}} \text{cl}(F_n)\right) = \left(\text{cl}\left(\bigcap_{n \in \mathbb{N}} \text{cl}(F_n)^c\right)\right)^c = X^c = \emptyset,$$

by Baire's category theorem, since each $\text{cl}(F_n)^c$ is open and dense. $\square$

Operators continuous in measure

Let $(X, \mathcal{M}, \mu)$ be a σ -finite measure space.

Definition

Let $T : L^p(X) \rightarrow L^p(X)$, with $1 \leq p < \infty$, be a linear operator. We say that T is continuous in measure if for every sequence $(f_k)_{k \in \mathbb{N}} \subseteq L^p(X)$ with $\|f_k\|_{L^p(X)} \xrightarrow{k \rightarrow \infty} 0$ and for every $\varepsilon > 0$, we have

$$\mu(\{x \in X : |T(f_k)(x)| > \varepsilon\}) \xrightarrow{k \rightarrow \infty} 0.$$

We say that a sequence of linear operators $(T_n)_{n \in \mathbb{N}}$ such that

$$T_n : L^p(X) \rightarrow L^p(X) \text{ for some } 1 \leq p < \infty$$

is continuous in measure, if each T_n is continuous in measure.

Remark

Remark

The assumption about continuity in measure is very mild and, in practice, can be easily verified. Indeed, in applications, we will usually work with a sequence of bounded linear operators $(T_n)_{n \in \mathbb{N}}$, which means that, for every $n \in \mathbb{N}$, there is $0 \leq C_n < \infty$ such that

$$\|T_n f\|_{L^p(X)} \leq C_n \|f\|_{L^p(X)} \quad \text{for all } f \in L^p(X).$$

Then, for every $\varepsilon > 0$, by Chebyshev's inequality, we have

$$\begin{aligned} \mu(\{x \in X : |T_n(f_k)(x)| > \varepsilon\}) &\leq \frac{1}{\varepsilon^p} \int_X |T_n(f_k)(x)|^p d\mu(x) \\ &= \frac{1}{\varepsilon^p} \|T_n(f_k)\|_{L^p}^p \leq \left(\frac{C_n}{\varepsilon}\right)^p \|f_k\|_{L^p}^p \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Thus, $(T_n)_{n \in \mathbb{N}}$ is continuous in measure.

Remark

Remark

If $T_n f(x)$ converges μ -almost everywhere on X , then the maximal function (maximal operator)

$$T_* f(x) = \sup_{n \in \mathbb{N}} |T_n f(x)|$$

is bounded μ -almost everywhere on X .

Now we will prove a certain uniform boundedness principle that asserts that the continuity in measure of a sequence of operators and the almost everywhere finiteness of the maximal operator imply continuity at 0 in measure of the maximal operator.

Banach continuity principle

Theorem (The Banach continuity principle)

Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Assume that:

- ① The sequence $(T_n)_{n \in \mathbb{N}}$ of linear operators $T_n : L^p(X) \rightarrow L^p(X)$ for some $1 \leq p < \infty$ is continuous in measure.
- ② The maximal operator $T_*f(x) = \sup_{n \in \mathbb{N}} |T_n f(x)| < \infty$ μ -a.e. on X .

Then there exists a decreasing function $C(\alpha)$ defined for all $\alpha > 0$ and such that $\lim_{\alpha \rightarrow \infty} C(\alpha) = 0$ and such that

$$\mu(\{x \in X : T_*f(x) > \alpha \|f\|_{L^p(X)}\}) \leq C(\alpha) \quad \text{for every } f \in L^p(X).$$

Proof. The idea is to define

$$C(\alpha) = \sup_{f \in L^p(X)} \mu(\{x \in X : T_*f(x) > \alpha \|f\|_{L^p(X)}\})$$

and show that $C(\alpha)$ is decreasing, which is clear, and $\lim_{\alpha \rightarrow \infty} C(\alpha) = 0$.

Proof

Given $\varepsilon > 0$, consider

$$F_n = \{f \in L^p(X) : \mu(\{x \in X : T_* f(x) > n\}) \leq \varepsilon\}.$$

Step 1: We show that F_n is closed in $L^p(X)$ for every $n \in \mathbb{N}$. To prove this, consider $f \notin F_n$, so then

$$\mu(\{x \in X : T_* f(x) > n\}) > \varepsilon.$$

It follows that there exists $N \in \mathbb{N}$ so that

$$\mu(\{x \in X : \sup_{1 \leq k \leq N} |T_k f(x)| > n\}) > \varepsilon.$$

Then there exists $\delta > 0$ such that

$$\mu(\{x \in X : \sup_{1 \leq k \leq N} |T_k f(x)| > n + \delta\}) > \varepsilon + \delta.$$

These two choices of N and δ are possible by the continuity of the measure μ .

Proof

Now, by the continuity in measure of the operators T_k , there exists a $\delta' > 0$ such that, for every

$$g \in B(f, \delta') = \{h \in L^p(X) : \|f - h\|_{L^p} < \delta'\},$$

we have for $1 \leq k \leq N$ that

$$\mu(\{x \in X : |T_k(f - g)(x)| > \delta\}) < \frac{\delta}{2^k}.$$

Let $Z = \bigcup_{k=1}^N \{x \in X : |T_k(f - g)(x)| > \delta\}$. Then $\mu(Z) \leq \delta$ and

$$\{x \in X : \sup_{1 \leq k \leq N} |T_k f(x)| > n + \delta\} \subseteq Z \cup \{x \in X : T_* g(x) > n\},$$

which implies that $\mu(\{x \in X : T_* g(x) > n\}) > \varepsilon$ for all $g \in B(f, \delta')$. Therefore, $L^p(X) \setminus F_n$ is open as desired.

Proof

Step 2: Since $T_*f(x) = \sup_{k \in \mathbb{N}} |T_k f(x)| < \infty$ μ -a.e. and $\mu(X) < \infty$, we have

$$L^p(X) = \bigcup_{n \in \mathbb{N}} F_n,$$

and we know the F_n are closed. Since $L^p(X)$ is complete, by the Baire category theorem, there is at least one $n \in \mathbb{N}$ with

$$\text{int}(F_n) \neq \emptyset.$$

Thus, there exists $f_0 \in F_n$ and $\delta > 0$ so that $f_0 + \delta g \in F_n$ for all $g \in L^p(X)$ with $\|g\|_{L^p} = 1$, and

$$\mu(\{x \in X : T_*(f_0 + \delta g)(x) > n\}) \leq \varepsilon.$$

Therefore,

$$\begin{aligned} \mu(\{x \in X : T_*g(x) \geq \frac{2n}{\delta}\}) &\leq \mu(\{x \in X : T_*(f_0 + \delta g)(x) > n\}) \\ &\quad + \mu(\{x \in X : T_*(f_0 - \delta g)(x) > n\}) \leq 2\varepsilon. \end{aligned}$$

Proof

Hence, for every $g \in L^p(X)$,

$$\mu(\{x \in X : T_*g(x) > \frac{2n}{\delta} \|g\|_{L^p}\}) \leq 2\varepsilon.$$

Setting

$$C(\alpha) = \sup_{f \in L^p(X)} \mu(\{x \in X : T_*f(x) > \alpha \|f\|_{L^p(X)}\}),$$

we have that $\lim_{\alpha \rightarrow \infty} C(\alpha) = 0$.



Hopf maximal inequality

Let $(X, \mathcal{M}, \mu)$ be a measure space and consider a linear operator $T : L^1(X) \rightarrow L^1(X)$ such that

- ① $f \geq 0$ μ -almost everywhere implies $Tf \geq 0$ μ -almost everywhere.
- ② $\|Tf\|_{L^1(X)} \leq \|f\|_{L^1(X)}$ for all $f \in L^1(X)$.
- ③ $\|Tf\|_{L^\infty(X)} \leq \|f\|_{L^\infty(X)}$ for all $f \in L^1(X) \cap L^\infty(X)$.

Remark

If $\mu(X) < \infty$, then (1) and $T\mathbb{1} = \mathbb{1}$ imply (3).

Lemma

Let $1 \leq p < \infty$, $0 \leq f \in L^p(X; \mathbb{R})$, and $\lambda > 0$. Then the following hold:

- ① $\mu(\{x \in X : f(x) > \lambda\}) \leq \lambda^{-p} \|f\|_{L^p}^p < \infty$.
- ② $(f - \lambda)^+ \in L^1(X) \cap L^p(X)$.
- ③ $Tf - \lambda \leq T(f - \lambda)^+$.

Proof

- ① This is just Chebyshev's inequality.
- ② Let $A = \{x \in X : f(x) > \lambda\}$, and note that

$$(f - \lambda)^+ = (f - \lambda)\mathbb{1}_A = f\mathbb{1}_A - \lambda\mathbb{1}_A.$$

Since $\mu(A) < \infty$, the claim holds.

- ③ Observe that

$$\begin{aligned} |f - (f - \lambda)^+| &= |f\mathbb{1}_A - (f - \lambda)^+\mathbb{1}_A + f\mathbb{1}_{A^c} - (f - \lambda)^+\mathbb{1}_{A^c}| \\ &= |\lambda\mathbb{1}_A + f\mathbb{1}_{A^c}| \leq \lambda. \end{aligned}$$

Therefore,

$$Tf - T(f - \lambda)^+ \leq |T(f - (f - \lambda)^+)| \leq \|f - (f - \lambda)^+\|_{L^\infty(X)} \leq \lambda.$$

This completes the proof of the Lemma. □

Notation

For $0 \leq f \in L^p(X)$ with $1 \leq p < \infty$ and $\lambda > 0$, we write

$$A_*^n f = \max_{1 \leq k \leq n} A_k f,$$

$$A_k f = \frac{1}{k} \sum_{m=0}^{k-1} T^m f,$$

$$S_k f = \sum_{m=0}^{k-1} T^m f,$$

$$M_\lambda^n f = \max_{1 \leq k \leq n} (S_k f - k\lambda).$$

Then the set

$$\begin{aligned} \{x \in X : A_*^n f(x) > \lambda\} &= \{x \in X : M_\lambda^n f(x) > 0\} \\ &\subseteq \bigcup_{k=1}^n \{x \in X : S_k f(x) > k\lambda\} \end{aligned}$$

has finite measure and $(M_\lambda^n f)^+ \in L^1(X) \cap L^p(X)$ by the previous lemma.

Hopf maximal inequality

Theorem (Hopf maximal theorem)

Let $(X, \mathcal{M}, \mu)$ be a measure space and consider a linear operator $T : L^1(X) \rightarrow L^1(X)$ such that

- ① $f \geq 0$ μ -almost everywhere implies $Tf \geq 0$ μ -almost everywhere.
- ② $\|Tf\|_{L^1(X)} \leq \|f\|_{L^1(X)}$, $f \in L^1(X)$.
- ③ $\|Tf\|_{L^\infty(X)} \leq \|f\|_{L^\infty(X)}$ for all $f \in L^1(X) \cap L^\infty(X)$.

Let $1 \leq p < \infty$ and $0 \leq f \in L^p(X)$. Then, for each $\lambda > 0$ and $n \in \mathbb{N}$, we have

$$\mu(\{x \in X : A_*^n f(x) > \lambda\}) \leq \frac{1}{\lambda} \int_{\{A_*^n f > \lambda\}} f d\mu.$$

Proof (after Garsia)

Proof. Take $k \in \{2, 3, \dots, n\}$ and note that

$$\begin{aligned} S_k f - k\lambda &= f - \lambda + TS_{k-1}f - (k-1)\lambda \\ &\leq f - \lambda + T(S_{k-1}f - (k-1)\lambda)^+ \\ &\leq f - \lambda + T(M_\lambda^n f)^+, \end{aligned}$$

(for $k = 1$, $S_k f - k\lambda = f - \lambda$). Taking the maximum over k , we obtain

$$M_\lambda^n f \leq f - \lambda + T(M_\lambda^n f)^+.$$

Now we see that

$$\begin{aligned} \int_X (M_\lambda^n f)^+ d\mu &= \int_{\{M_\lambda^n f > 0\}} M_\lambda^n f d\mu \\ &\leq \int_{\{M_\lambda^n f > 0\}} (f - \lambda) d\mu + \int_X T(M_\lambda^n f)^+ d\mu \\ &\leq \int_{\{M_\lambda^n f > 0\}} (f - \lambda) d\mu + \int_X (M_\lambda^n f)^+ d\mu. \end{aligned}$$

Proof (after Garsia)

After subtracting from both sides, we have

$$0 \leq \int_{\{M_\lambda^n f > 0\}} (f - \lambda) d\mu,$$

and this means that

$$\lambda \mu(\{x \in X : M_\lambda^n f(x) > 0\}) \leq \int_{\{M_\lambda^n f > 0\}} f d\mu,$$

and

$$\mu(\{x \in X : M_\lambda^n f(x) > 0\}) \leq \frac{1}{\lambda} \int_{\{M_\lambda^n f > 0\}} f d\mu,$$

and

$$\mu(\{x \in X : A_*^n f(x) > \lambda\}) \leq \frac{1}{\lambda} \int_{\{A_*^n f > \lambda\}} f d\mu.$$

This completes the proof. □

Hopf maximal inequality for complex-valued functions

Remark

We can extend the result to the case where $f \in L^p(X)$ is a complex-valued function and T is positive, i.e., $|Tf| \leq T|f|$. Indeed, $|Tf| = \alpha Tf$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Thus, $|Tf| = \operatorname{Re}(\alpha Tf) = \operatorname{Re}(T(\alpha f))$. But $Tf \in \mathbb{R}$ for every $f : X \rightarrow \mathbb{R}$. Since T is positive and $f = f^+ - f^-$, we have

$$|Tf| = \operatorname{Re}(T(\alpha f)) = \operatorname{Re}(T(\operatorname{Re}(\alpha f)) + iT(\operatorname{Im}(\alpha f))) = T(\operatorname{Re}(\alpha f)).$$

Therefore,

$$|Tf| = T(\operatorname{Re}(\alpha f)) \leq T|\operatorname{Re}(\alpha f)| \leq T|f|$$

since $|\operatorname{Re}(f)| \leq |f|$. Thus,

$$\mu(\{x \in X : A_*^n |f|(x) > \lambda\}) \leq \frac{1}{\lambda} \int_{\{A_*^n |f| > \lambda\}} |f| d\mu.$$

Hopf maximal inequality for L^p spaces

Corollary

Under the assumptions of the previous theorem, we have for all $f \in L^p(X)$ with $1 \leq p \leq \infty$ that

$$\|A_*^n f\|_{L^p(X)} \leq \frac{p}{p-1} \|f\|_{L^p(X)}.$$

Proof. First note that $\|A_*^n f\|_{L^p(X)} \leq \|A_*^n |f|\|_{L^p(X)}$.

$$\begin{aligned} \|A_*^n |f|\|_{L^p(X)}^p &= p \int_0^\infty \lambda^{p-1} \mu(\{x \in X : A_*^n |f|(x) > \lambda\}) d\lambda \\ &\leq p \int_0^\infty \lambda^{p-2} \int_{\{A_*^n |f| > \lambda\}} |f| d\mu d\lambda \\ &= p \int_X |f| \left(\int_0^{A_*^n |f|} \lambda^{p-2} d\lambda \right) d\mu \\ &= \frac{p}{p-1} \int_X |f| (A_*^n |f|)^{p-1} d\mu \leq \frac{p}{p-1} \|f\|_{L^p} \|A_*^n |f|\|_{L^p}^{p-1}. \quad \square \end{aligned}$$

Birkhoff's ergodic theorem

Theorem (Birkhoff's ergodic theorem)

Let $(X, \mathcal{M}, \mu, T)$ be a σ -finite measure-preserving system. For $1 \leq p \leq \infty$ and $f \in L^p(X)$ define the ergodic average by

$$A_N f(x) = \frac{1}{N} \sum_{m=0}^{N-1} f(T^m x) \quad \text{for } x \in X.$$

If $1 \leq p < \infty$ and $f \in L^p(X)$ then there exists a T -invariant function $f^* \in L^p(X)$ such that

$$\lim_{N \rightarrow \infty} A_N f(x) = f^*(x) \quad \mu\text{-a.e. on } X.$$

If additionally, $0 < \mu(X) < \infty$ and T is ergodic, then

$$\lim_{N \rightarrow \infty} A_N f(x) = \frac{1}{\mu(X)} \int_X f(x) d\mu(x) \quad \mu\text{-a.e. on } X.$$

Proof

Proof. We shall use the two-step procedure to establish pointwise convergence. Let

$$A_*f(x) = \sup_{N \in \mathbb{N}} |A_Nf(x)| \quad \text{for } x \in X.$$

be the maximal function corresponding to the ergodic averages A_Nf .

Step 1. By Hopf's maximal inequality we may derive the following two maximal estimates for A_*f for any $f \in L^p(X)$ with $1 \leq p \leq \infty$, i.e.

- If $p = 1$ we have weak-type $(1, 1)$ maximal inequality

$$\mu(\{x \in X : A_*f(x) > \lambda\}) \leq \frac{1}{\lambda} \|f\|_{L^1(X)} \quad \text{for } \lambda > 0.$$

- If $1 < p \leq \infty$ we have strong-type (p, p) maximal inequality

$$\|A_*f\|_{L^p(X)} \leq \frac{p}{p-1} \|f\|_{L^p(X)}.$$

This ensures that the set of $L^p(X)$ functions for which we have pointwise convergence for A_Nf is closed in $L^p(X)$.

Proof

Step 2. In view of the previous step we have to find a dense class of functions $\mathcal{D} \subseteq L^p(X)$ such that $A_N f$ converges μ -a.e. on X for $f \in \mathcal{D}$. We now distinguish two cases.

- Suppose that $p = 2$ and we use Riesz's decomposition. Let

$$I_T = \{f \in L^2(X) : f \circ T = f\},$$

$$J_T = \{g - g \circ T : g \in L^2(X) \cap L^\infty(X)\}.$$

Then one sees that

$$\overline{I_T \oplus J_T}^{\|\cdot\|_{L^2(X)}} = L^2(X),$$

which means that $\mathcal{D} = I_T \oplus J_T$ is dense in $L^2(X)$. If $f \in I_T$, then clearly $A_N f = f$ μ -a.e. on X , and we have

$$\lim_{N \rightarrow \infty} A_N f(x) = f(x) \quad \mu\text{-a.e. on } X.$$

Proof

If $f \in J_T$, then $f = g \circ T - g$ and by telescoping we deduce that

$$\begin{aligned} |A_N f(x)| &= \left| \frac{1}{N} \sum_{m=0}^{N-1} (g(T^{m+1}x) - g(T^m x)) \right| \\ &= \frac{1}{N} |g(T^N x) - g(x)| \leq \frac{2}{N} \|g\|_{L^\infty(X)} \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

μ -a.e. on X . Thus we have established pointwise almost everywhere convergence of $A_N f$ on a dense class $\mathcal{D} = I_T \oplus J_T$ in $L^2(X)$, which combined with the maximal estimate for $p = 2$ gives pointwise almost everywhere convergence for all $f \in L^2(X)$.

- Suppose that $p \neq 2$. Then it suffices to observe that $L^2(X) \cap L^p(X)$ is dense in $L^p(X)$ and $A_N f$ converges pointwise almost everywhere for any $f \in L^2(X) \cap L^p(X)$. Now combining this fact with the maximal estimates for all $1 \leq p < \infty$ we deduce pointwise almost everywhere convergence of $A_N f$ for all $f \in L^p(X)$ as desired.

Proof

- We now show that the limit

$$\lim_{N \rightarrow \infty} A_N f(x) = f^*(x)$$

is T -invariant. Indeed, note that

$$\begin{aligned} f^*(Tx) &= \lim_{N \rightarrow \infty} A_N f(Tx) \\ &= \lim_{N \rightarrow \infty} \left(\frac{N+1}{N} A_{N+1} f(x) - \frac{1}{N} f(x) \right) = f^*(x) \end{aligned}$$

μ -a.e. on X , which shows that f^* is T -invariant.

- Suppose now that $0 < \mu(X) < \infty$ and T is ergodic. We may assume, without loss of generality, that $\mu(X) = 1$.
- Since $f^*(Tx) = f^*(x)$ μ -a.e. on X , and T is ergodic, we deduce that f^* is constant μ -a.e. on X .

Proof

- We now prove that

$$f^*(x) = \int_X f(x) d\mu(x) \quad \mu\text{-a.e. on } X.$$

Indeed, if $1 < p \leq \infty$ we have strong-type (p, p) maximal inequality

$$\|A_* f\|_{L^p(X)} \leq \frac{p}{p-1} \|f\|_{L^p(X)},$$

then we can use the **(DCT)** to deduce that

$$f^*(x) = \int_X f^*(x) d\mu(x) = \lim_{N \rightarrow \infty} \int_X A_N f(x) d\mu(x) = \int_X f(x) d\mu(x)$$

μ -a.e. on X , since f^* is constant μ -a.e., and μ is T -invariant.

- For $p = 1$ we have to be more careful.

Proof

- For $p = 1$ we have to be more careful.
- Let $\varepsilon > 0$ and split the function $f = g + h$, where $g \in L^2(X)$ and $h \in L^1(X)$ with

$$\|h\|_{L^1(X)} < \varepsilon.$$

Then

$$\begin{aligned} \left| f^*(x) - \int_X f(t) \mu(dt) \right| &= \left| h^*(x) - \int_X h(t) \mu(dt) \right| \\ &\leq 2\|h\|_{L^1(X)} < 2\varepsilon. \end{aligned}$$

This completes the proof of Birkhoff's theorem. □

Remark

If $T : X \rightarrow X$ is ergodic on an infinite σ -finite measure space $(X, \mathcal{M}, \mu)$, then

$$\lim_{N \rightarrow \infty} A_N f(x) = 0 \quad \mu\text{-a.e. on } X.$$

Equidistribution and Weyl's criterion

A sequence $(a_k)_{k \in \mathbb{N}} \subseteq [0, 1]$ is called equidistributed if for every continuous function $f : [0, 1] \rightarrow \mathbb{C}$ we have that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(a_k) = \int_0^1 f(x) dx.$$

Theorem (Weyl's theorem)

The following statements are equivalent:

- (a) *The sequence $(a_k)_{k \in \mathbb{N}} \subseteq [0, 1]$ is equidistributed.*
- (b) *For every $m \in \mathbb{Z} \setminus \{0\}$ we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e^{2\pi i m a_k} = 0.$$

- (c) *For any $[a, b] \subset [0, 1]$ we have*

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : a_n \in [a, b]\}}{N} = b - a.$$

Classical equidistribution problems

- Around 1910, Weyl (also independently Bohl and Sierpiński) proved that for every irrational $\theta \in \mathbb{R}$, any $[a, b] \subset [0, 1)$ one has

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : \{\theta n\} \in [a, b]\}}{N} = b - a.$$

- In 1916, Weyl showed that $(\{P(n)\})_{n \in \mathbb{N}}$ is equidistributed for every polynomial $P : \mathbb{R} \rightarrow \mathbb{R}$ having at least one irrational coefficient.
- In 1933 Khinchin had the great insight to see how to generalize the classical equidistribution result by using Birkhoff's ergodic theorem and proved that for any irrational $\theta \in \mathbb{R}$, for any Lebesgue measurable set $E \subseteq [0, 1)$, and for almost every $x \in \mathbb{R}$, one has

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N : \{x + \theta n\} \in E\}}{N} = |E|,$$

Consequences of the Birkhoff ergodic theorem

- **(Borel's Theorem on Normal Numbers).** Almost all numbers in $[0, 1)$ are normal to base 2, i.e. for a.e. $x \in [0, 1)$ the frequency of 1's in the binary expansion of x is $1/2$.
- **(Frequency of the natural number k in the partial quotients).** For almost every real number $x \in (0, 1)$, the digit k appears in the continued fraction expansion $x = [a_1, a_2, \dots]$ with density

$$(\log 2)^{-1} \log \left(1 + \frac{1}{k^2 + 2k} \right).$$

- **(Strong law of large numbers).** If $X_1, X_2, \dots$ is an infinite sequence of i.i.d. integrable random variables with mean μ , then

$$\lim_{N \rightarrow \infty} \frac{1}{N} (X_1 + \dots + X_N) = \mu.$$

- **(Kac theorem).** Let $(X, \mathcal{M}, \mu, T)$ be a probability measure-preserving system $\mu(X) = 1$ and assume that T is ergodic. Then for any $A \in \mathcal{M}$ with $\mu(A) > 0$ the expected return time to A is $\mu(A)^{-1}$, equivalently $\int_A \inf\{n \in \mathbb{N} : T^n(x) \in A\} d\mu(x) = 1$.