On pointwise convergence problems, part II

Mariusz Mirek

IAS Princeton & Rutgers University

Uniwersytet Jagielloński
Kraków May 17, 2023

Supported by the NSF grant DMS-2154712,
and the CAREER grant DMS-2236493.

Khinchin's equidistribution theorem

Theorem (Birkhoff ergodic theorem, (1931))

Let $(X, \mathcal{B}(X), \mu, T)$ be a σ-finite measure preserving system. For every $1 \leq p < \infty$ and every $f \in L^p(X)$ there exists $f^* \in L^p(X)$ such that

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) \underset{N\to\infty}{\longrightarrow} f^*(x)$$

- 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

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

- A famous Bellow problem from the early 1980's asks whether the same conclusion holds in Khinchin's result if we replace n with any polynomial $P(n)$ with integer coefficients.

Bourgain's pointwise ergodic theorem

In the early 1980's Bellow and Furstenberg asked independently about the poinwise convergence for polynomial ergodic averages

$$A_N^P f(x) := \frac{1}{N}\sum_{n=1}^{N} f(T^{P(n)}x) \quad \text{for} \quad x \in X,$$

where $P \in \mathbb{Z}[\mathrm{n}]$ is a polynomial of degree > 1.

- Furstenberg was motivated by the result of Sárközy: $S \subseteq \mathbb{Z}$ has positive upper Banach density, then there are $x, n \in \mathbb{N}$ such that $x, x + n^2 \in S$.
- Furstenberg proved norm convergence for $A_N^P f$ and deduced the polynomial Poincaré recurrence theorem: if $\mu(X) < \infty$ and $E \in \mathcal{B}(X)$ with $\mu(E) > 0$, then $\mu(E \cap T^{-P(n)}[E]) > 0$ for some $n \in \mathbb{N}$.

Bellow and Furstenberg question was very hard. Even for $P(n) = n^2$, since $(n+1)^2 - n^2 = 2n + 1$. For overcoming this problem, Bourgain used the ideas from the circle method of Hardy and Littlewood to show:

- $L^p(X)$ boundedness of the maximal function for any $1 < p \leq \infty$.
- Given an increasing sequence $(N_j : j \in \mathbb{N})$, for each $J \in \mathbb{N}$ one has

$$\Big(\sum_{j=0}^{J} \big\| \sup_{N_j \leq N < N_{j+1}} \big|A_N^P f - A_{N_j}^P f\big| \big\|_{L^2(X)}^2\Big)^{1/2} \leq o(J^{1/2}) \|f\|_{L^2(X)}.$$

Furstenberg–Bergelson–Leibman conjecture

One of the central open problems in pointwise ergodic theory (from the mid 1980's) is a conjecture of Furstenberg–Bergelson–Leibman:

Theorem (Furstenberg–Bergelson–Leibman conjecture)

Let $\mathbb{G}$ be a nilpotent group of measure preserving transformations of a probability space $(X, \mathcal{B}(X), \mu)$. Let $P_{j,i} \in \mathbb{Z}[\mathrm{n}]$ be polynomials and $T_1, \ldots, T_d \in \mathbb{G}$ and $f_1, \ldots, f_m \in L^\infty(X)$. Does the limit of the averages

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T_1^{P_{1,1}(n)} \cdots T_d^{P_{1,d}(n)} x) \cdot \ldots \cdot f_m(T_1^{P_{m,1}(n)} \cdots T_d^{P_{m,d}(n)} x) \tag{1}$$

exist μ-almost everywhere on X as $N \to \infty$?

- The norm convergence in $L^2(X)$ for the averages (1) was established in the nilpotent setting by M. Walsh in 2012 .
- Bergelson and Leibman showed that $L^2(X)$ norm convergence for (1) may fail if $\mathbb{G}$ is a solvable group.
- The nilpotent setting is probably the most general setting where the conjecture of Furstenberg–Bergelson–Leibman might be true.

Recent contribution to the nilpotent setting

Linear and nilpotent variant of the Furstenberg–Bergelson–Leibman problem can be summarize as follows:

Theorem (M., Ionescu, Magyar, and Szarek (2021))

Let $(X, \mathcal{B}(X), \mu)$ be a σ-finite space and let $T_1, \ldots, T_d : X \to X$ be a family of invertible and measure preserving transformations satisfying

$$[[T_i, T_j], T_k] = \mathrm{Id} \qquad \textit{for all} \quad 1 \le i \le j \le k \le d.$$

Then for every polynomials $P_1, \ldots, P_d \in \mathbb{Z}[\mathrm{n}]$ and every $f \in L^p(X)$ with $1 < p < \infty$ the averages

$$\frac{1}{N} \sum_{n=1}^{N} f(T_1^{P_1(n)} \cdots T_d^{P_d(n)} x)$$

converge for μ-almost every $x \in X$ and in $L^p(X)$ norm as $N \to \infty$.

- One can think that $T_1, \ldots, T_d$ belong to a nilpotent group of step two of measure preserving mappings of a σ-finite space $(X, \mathcal{B}(X), \mu)$.

Recent contribution to the bilinear setting

After Bourgain's pointwise bilinear ergodic theorem for

$$A_N^{an,bn}(f,g)(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^{an}x)g(T^{bn}x) \qquad a,b \in \mathbb{Z}$$

jointly with Ben Kruse and Terry Tao we established the following theorem.

Theorem (Krause, M., and Tao, (2020))

Let $(X, \mathcal{B}(X), \mu, T)$ be an invertibe σ-finite measure-preserving system, let $P \in \mathbb{Z}[\mathrm{n}]$ with $\deg(P) \geq 2$, and let $f \in L^{p_1}(X)$ and $g \in L^{p_2}(X)$ for some $p_1, p_2 \in (1, \infty)$ with $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p} \leq 1$. Then the Furstenberg–Weiss averages

$$A_N^{\mathrm{n},P(\mathrm{n})}(f,g)(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^{n}x)g(T^{P(n)}x)$$

converge for μ-almost every $x \in X$ and in $L^p(X)$ norm as $N \to \infty$.

Corollary

For any irrational $\theta \in \mathbb{R}$, for any Lebesgue measurable sets $E, F \subseteq [0,1)$, and for almost every $x \in \mathbb{R}$, one has

$$\lim_{N\to\infty} \frac{\#\{1 \leq n \leq N : \{x + \theta n\} \in E, \{x + \theta P(n)\} \in F\}}{N} = |E||F|.$$

Key ideas

The proof of pointwise convergence for

$$A_N^{\mathrm{n},P(\mathrm{n})}(f,g)(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^n x) g(T^{P(n)} x)$$

is quite intricate, and relies on several deep results in the literature:

- the Ionescu–Wainger multiplier theorem (discrete Littlewood–Paley theory and paraproduct theory) — (HA)&(NT);
- The circle method of Hardy and Littlewood — (NT);
- the inverse theory of Peluse and Prendeville — (CO)&(NT);
- Hahn–Banach separation theorem — (FA);
- L^p-improving estimates of Han–Kovač–Lacey–Madrid–Yang (derived from the Vinogradov mean value theorem) — (HA)&(NT);
- Rademacher–Menshov argument combined with Khinchine's inequality — (HA)&(FA)&(PR);
- $L^p(\mathbb{R})$ bounds for a shifted square function — (HA);
- bounded metric entropy argument from Banach space theory— (CO)&(FA)&(PR);
- van der Corput type estimates in the p-adic fields — (HA)&(NT).

Inverse theorem of Peluse and Prendiville

An important ingredient in the proof is the inverse theorem of Peluse, which can be thought of as a counterpart of Weyl's inequality:

Theorem (Peluse, (2019/2020))

Let $m \geq 2$ and $P_1, \ldots, P_m \in \mathbb{Z}[\mathrm{n}]$ each having zero constant term such that $\deg P_1 < \ldots < \deg P_m$. *Let $N \in \mathbb{N}$ and $\delta \in (0,1)$ and assume that functions $f_0, f_1, \ldots, f_m : \mathbb{Z} \to \mathbb{C}$ are supported on $[-N_0, N_0]$ for some $N_0 \simeq N^{\deg P_m}$, and $\|f_0\|_{L^\infty(\mathbb{Z})}, \|f_1\|_{L^\infty(\mathbb{Z})}, \ldots, \|f_m\|_{L^\infty(\mathbb{Z})} \leq 1$, and suppose that*

$$\Big\| \frac{1}{N} \sum_{n=1}^{N} f_0(x) f_1(x - P_1(n)) \cdots f_m(x - P_m(n)) \Big\|_{L^1_x(\mathbb{Z})} \geq \delta N^{\deg P_m}.$$

Then there are $1 \leq q \lesssim \delta^{-O(1)}$ and $\delta^{O(1)} N^{\deg P_1} \lesssim M \leq N^{\deg P_1}$ such that

$$\Big\| \frac{1}{M} \sum_{y=1}^{M} f_1(x + qy) \Big\|_{L^1_x(\mathbb{Z})} \gtrsim \delta^{O(1)} N^{\deg P_m}$$

provided that $N \gtrsim \delta^{-O(1)}$.

Bilinear Weyl's inequality

In the multilinear theory Weyl's inequality is inefficient, and inability to invoke Plancherel's theorem forced us to proceed differently. We proved the following bound on minor arcs:

Theorem (M., Krause, and Tao)

Let $P \in \mathbb{Z}[\mathrm{n}]$, $\deg P \geq 2$ *and* $P(0) = 0$. *Let* $\mathcal{M}_{\leq l, \leq k} = \bigsqcup_{a/q \in \Sigma_{\leq l}} [\frac{a}{q} - 2^{-k}, \frac{a}{q} + 2^{-k}]$ *be a family of major arcs corresponding to* $\Sigma_{\leq l}$. *If functions* $f, g \in \ell^2(\mathbb{Z})$ *and*

1. *either* $\hat{f}$ *vanishes on* $\mathcal{M}_{\leq u_k, \leq -k+u_k}$,
2. *or* $\hat{g}$ *vanishes on* $\mathcal{M}_{\leq u_k, \leq -dk+u_k}$, *then we have*

$$\Big\| 2^{-k} \sum_{n=1}^{2^k} f(x-n) g(x-P(n)) \Big\|_{\ell^1(\mathbb{Z})} \lesssim k^{-10} \|f\|_{\ell^2(\mathbb{Z})} \|g\|_{\ell^2(\mathbb{Z})}.$$

This theorem is the core of our argument and its proof is quite intricate, and relies on several deep results in the literature, including:

- the Ionescu–Wainger multiplier theorem (discrete Littlewood–Paley theory);
- the inverse theory of Peluse and Prendeville;
- Hahn–Banach separation theorem;
- L^p-improving estimates of Han–Kovač–Lacey–Madrid–Yang (derived from the Vinogradov mean value theorem).

Quantitative polynomial Szemerédi's

Let $r_{P_1,\ldots,P_m}(N)$ denote the size of the largest subset of $\{1,\ldots,N\}$ containing no configuration of the form $x, x+P_1(n),\ldots,x+P_m(n)$ with $n \neq 0$.

- Berglson and Leibman showed proving polynomial multiple recurrence theorem that
$$r_{P_1,\ldots,P_m}(N) = o_{P_1,\ldots,P_m}(N),$$
whenever $P_1,\ldots,P_m \in \mathbb{Z}[\mathrm{n}]$ and each having zero constant term.

Theorem (Gowers (2001), higher order Fourier analysis)

If $P_1(n) = n,\ldots,P_m(n) = (m-1)n$ *for every* $m \in \mathbb{N}$ *then there is* $\gamma_m > 0$ *such that*
$$r_{P_1,\ldots,P_m}(N) \lesssim \frac{N}{(\log\log N)^{\gamma_m}}.$$

- No bounds were known in general for the polynomial Szemerédi's theorem until a series of recent papers of Peluse and Prendiville.
- Peluse showed that there is a constant $\gamma_{P_1,\ldots,P_m} > 0$ such that
$$r_{P_1,\ldots,P_m}(N) \lesssim_{P_1,\ldots,P_m} \frac{N}{(\log\log N)^{\gamma_{P_1,\ldots,P_m}}}$$
answering a question posed by Gowers.

Commutative Furstenberg–Bergelson–Leibman conjecture

Ongoing project (Krause, M., Peluse, and Wright (2021))

Let $(X, \mathcal{B}(X), \mu)$ be a probability space equipped with commuting invertible measure-preserving maps $T_1, \ldots, T_k : X \to X$. Consider $P_1, \ldots, P_k \in \mathbb{Z}[\mathrm{n}]$ with distinct degrees and $f_1, \ldots, f_k \in L^\infty(X)$. It is expected that the averages

$$A_N^{P_1,\ldots,P_k}(f_1, \ldots, f_k)(x) = \frac{1}{N} \sum_{n=1}^{N} f_1(T_1^{P_1(n)} x) \ldots f_k(T_k^{P_k(n)} x)$$

converge for μ-almost every $x \in X$.

- There is some hope in the case when $T_1 = \ldots = T_k = T$.
- We also have made some progress for the following averages

$$\frac{1}{N} \sum_{n=1}^{N} f(T_1^n x) g(T_2^{n^2} x)$$

that correspond to the "squorners": $(x, y), (x+n, y), (x, y+n^2) \in \mathbb{Z}^2$, as well as to the following equidistribution result for $\alpha, \beta \in \mathbb{R} \setminus \mathbb{Q}$:

$$\lim_{N\to\infty} \frac{\#\{1 \le n \le N : \{x + \alpha n\} \in E, \{x + \beta n^2\} \in F\}}{N} = |E||F|.$$

Riesz decomposition

- Let $U_T : L^2(X) \to L^2(X)$ be the operator associated with T defined by

$$U_T(x) = f \circ T(x) = f(Tx).$$

It is easy to see that for any $f_1, f_2 \in L^2(X, \mu)$

$$\langle U_T f_1, U_T f_2 \rangle = \langle f_1, f_2 \rangle,$$

we have hence U_T is an isometry on $L^2(X, \mu)$.

- Let us define

$$\mathbf{I}_T = \{f \in L^2(X) : f \circ T = f\}.$$

Lemma

For every σ-finite measure-preserving system $(X, \mathcal{B}(X), \mu)$ one has

$$L^2(X) = \mathbf{I}_T \oplus \overline{\mathbf{J}_T},$$

where

$$\mathbf{J}_T = \{g \circ T - g : g \in L^2(X)\},$$

and $\overline{\mathbf{J}_T}$ is the closure of $\mathbf{J}_T$ in $L^2(X)$.

Proof of Riesz decomposition

Proof.

The proof will be completed if we show that $\mathbf{I}_T = \mathbf{J}_T^{\perp}$.

- For the inclusion '$\subseteq$' observe that if $U_T f = f$ then we have

$$\langle f, U_T g - g\rangle = \langle U_T f, U_T g\rangle - \langle f, g\rangle = 0,$$

  hence $\mathbf{I}_T \subseteq \mathbf{J}_T^{\perp}$.
- For the opposite inclusion '$\supseteq$' note that if $f \in \mathbf{J}_T^{\perp}$ then for all $g \in L^2(X)$ we have

$$\langle U_T g, f\rangle = \langle g, f\rangle,$$

  hence $U_T^* f = f$. Therefore, $f = U_T f$ since

$$\begin{aligned} \|U_T f - f\|_{L^2(X)}^2 &= \|U_T f\|_{L^2(X)}^2 - \langle U_T f, f\rangle - \langle f, U_T f\rangle + \|f\|_{L^2(X)}^2 \\ &= 2\|f\|_{L^2(X)}^2 - \langle U_T^* f, f\rangle - \langle f, U_T^* f, \rangle = 0. \end{aligned}$$

This completes the proof of the lemma. □

von Neumann's ergodic theorem

Now we are able to prove von Neumann's mean ergodic theorem.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system then for every $f \in L^2(X)$ the averages

$$A_N f(x) = \frac{1}{N} \sum_{n=1}^{N} f(T^n x)$$

converge in $L^2(X)$ to Pf the orthogonal projection in $L^2(X)$ of f onto the space

$$\mathbf{I}_T = \{f \in L^2(X) : f \circ T = f\}.$$

Proof.

In view of Riesz decomposition it suffices to prove von Neumann's theorem for any function $f = f_1 + f_2$ where $f_1 \in \mathbf{I}_T$ and $f_2 \in \mathbf{J}_T$.

- For $f_1 \in \mathbf{I}_T$ our result is obvious since $A_N f_1 = f_1$ for every $N \in \mathbb{N}$.

□

Proof von Neumann's ergodic theorem

- If $f_2 \in \mathbf{J}_T$ then $f_2 = g \circ T - g$ for some $g \in L^2(X)$ and

$$\begin{aligned}\|A_N f_2\|_{L^2(X)} &= \left\| \frac{1}{N} \sum_{n=1}^{N} \left(g \circ T^{n+1} - g \circ T^{n} \right) \right\|_{L^2(X)} \\ &= \frac{1}{N} \left\| g \circ T^{N+1} - g \circ T \right\|_{L^2(X)} \le \frac{2}{N} \|g\|_{L^2(X)} \xrightarrow[N\to\infty]{} 0.\end{aligned}$$

The proof of Theorem 10 is completed. □

Question

- What about pointwise almost everywhere convergence for $A_N f$ whenever $f \in L^2(X)$?
- By a general theorem of Riesz we know that norm convergence (or even convergence in measure) implies convergence pointwise almost everywhere of $A_{N_k} f$ for certain subsequence $(N_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$.

Hardy–Littlewood maximal inequality on $\mathbb{Z}$

Theorem

For a finitely supported function $f : \mathbb{Z} \to \mathbb{C}$ *and for every* $N \in \mathbb{N}$ *define*

$$M_N f(x) = \frac{1}{N} \sum_{k=1}^{N} f(x-k),$$

which is $A_{N,\mathbb{Z},S} f$ *with* $X = \mathbb{Z}$ *and the shift operator* $S(x) = x - 1$. *Let*

$$E_\lambda = \{x \in \mathbb{Z} : \sup_{N \in \mathbb{N}} |M_N f(x)| > \lambda\}, \qquad \lambda > 0.$$

Then there is $C > 0$ *such that for every* $\lambda > 0$ *we have*

$$|E_\lambda| \leq \frac{C}{\lambda} \sum_{x \in E_\lambda} |f(x)| \leq \frac{C}{\lambda} \|f\|_{\ell^1(\mathbb{Z})}, \tag{2}$$

and if $1 < p \leq \infty$ *then*

$$\| \sup_{N \in \mathbb{N}} |M_N f| \|_{\ell^p(\mathbb{Z})} \leq \frac{Cp}{p-1} \|f\|_{\ell^p(\mathbb{Z})}. \tag{3}$$

Proof of the Hardy–Littlewood maximal inequality on $\mathbb{Z}$

To prove (2) we will use Vitali type argument. Let $f \geq 0$ and define

$$E_\lambda^N = \{x \in \mathbb{Z} \cap [-N, N] : \sup_{N \in \mathbb{N}} M_N f(x) > \lambda\}.$$

- For every $x \in E_\lambda^N$ there is $N_x \in \mathbb{N}$ such that

$$M_{N_x} f(x) = \frac{1}{N_x} \sum_{k=1}^{N_x} f(x-k) > \lambda.$$

- Then we see that

$$E_\lambda^N \subseteq \bigcup_{x \in E_\lambda^N} B_x,$$

  where $B_x = x + (-N_x, N_x) \subseteq E_\lambda^N$.

- Note that it is easy to find a finite sub-collection of disjoint intervals $\{B_{x_1}, B_{x_2}, \ldots, B_{x_J}\}$, such that $N_{x_1} \geq N_{x_2} \geq \ldots \geq N_{x_J}$ and

$$\bigcup_{n \in E_\lambda^N} B_n \subseteq \bigcup_{n=1}^{J} 3B_{x_n} = \bigcup_{n=1}^{J} (x_n - 3N_{x_n}, x_n + 3N_{x_n}).$$

Proof of the Hardy–Littlewood maximal inequality on $\mathbb{Z}$

- The
$$|E_\lambda^N| \leq \Big| \bigcup_{n\in E_\lambda^N} B_n \Big| \leq 3 \Big| \bigcup_{n=1}^{J} B_{x_n} \Big| = 3 \sum_{n=1}^{J} |B_{x_n}|.$$
- Finally, we obtain
$$\begin{aligned} |E_\lambda^N| &\leq 3 \sum_{n=1}^{J} |B_{x_n}| \\ &\leq \frac{9}{\lambda} \sum_{n=1}^{J} \sum_{k\in(-N_{x_n}, N_{x_n})} f(x_n - k) \\ &\leq \frac{9}{\lambda} \sum_{x\in E_\lambda^N} f(x). \end{aligned}$$

This completes the proof of inequality (2).

Proof of the Hardy–Littlewood maximal inequality on $\mathbb{Z}$

In the proof of inequality (3) we will use (2). Indeed, by (2), Fubini theorem and Hölder's inequality we have

$$\begin{aligned}
\|\sup_{N\in\mathbb{N}} M_N f\|_{\ell^p(\mathbb{Z})}^p &= p\int_0^\infty \lambda^{p-1}|\{x\in\mathbb{Z} : \sup_{N\in\mathbb{N}}|M_N f(x)|>\lambda\}|d\lambda \\
&\le 9p\int_0^\infty \lambda^{p-2}\sum_{x\in E_\lambda}|f(x)|d\lambda \\
&= 9p\sum_{x\in\mathbb{Z}}|f(x)|\int_0^{\sup_{N\in\mathbb{N}}|M_N f(x)|}\lambda^{p-2}d\lambda \\
&= \frac{9p}{p-1}\sum_{x\in\mathbb{Z}}|f(x)|\sup_{N\in\mathbb{N}}|M_N f(x)|^{p-1} \\
&\le \frac{9p}{p-1}\|f\|_{\ell^p(\mathbb{Z})}\|\sup_{N\in\mathbb{N}} M_N f(x)\|_{\ell^p(\mathbb{Z})}^{p-1},
\end{aligned}$$

and the proof is finished. □

Calderón transference principle

Theorem

Assume that $B \subseteq \mathbb{Z}$ such that $|B| = \infty$. Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system with the averages

$$\mathcal{A}_{N;X,T} f(x) = \frac{1}{|B \cap [0,N]|} \sum_{n \in B \cap [0,N]} f(T^n x).$$

Let $\mathcal{M}_N$ denote $\mathcal{A}_{N;\mathbb{Z},S}$ on $(\mathbb{Z}, \mathbf{P}(\mathbb{Z}), |\cdot|, S)$ with $S(x) = x + 1$, i.e.

$$\mathcal{M}_N f(x) = \frac{1}{|B \cap [0,N]|} \sum_{n \in B \cap [0,N]} f(x+n).$$

If for some $p \geq 1$ there is $C_p > 0$ such that

$$\| \sup_{N \in \mathbb{N}} |\mathcal{M}_N F| \|_{\ell^p(\mathbb{Z})} \leq C_p \|f\|_{\ell^p(\mathbb{Z})}, \qquad F \in \ell^p(\mathbb{Z}), \tag{4}$$

then

$$\| \sup_{N \in \mathbb{N}} |\mathcal{A}_{N;X,T} f| \|_{L^p(X)} \leq C_p \|f\|_{L^p(X)}, \qquad f \in L^p(X). \tag{5}$$

Proof of the Calderón transference principle

Assume that $p \geq 1$. Let $J, R \in \mathbb{N}, f \in L^p(X)$ and define

$$F(j) = \begin{cases} f(T^j x) & 0 \leq j \leq J, \\ 0 & \text{otherwise}, \end{cases}$$

For a fixed $N \in \mathbb{N}$ such that $1 \leq N \leq R$ and every $0 \leq j \leq J - R$ we have

$$\begin{aligned} \mathcal{M}_N F(j) &= \frac{1}{|B \cap [0,N]|} \sum_{k \in B \cap [0,N]} F(j+k) \\ &= \frac{1}{|B \cap [0,N]|} \sum_{k \in B \cap [0,N]} f(T^{j+k} x) \\ &= \mathcal{A}_{N;X,T} f(T^j x). \end{aligned}$$

Thus for $1 \leq N \leq R$ we have

$$\begin{aligned} \sum_{j=0}^{J-R} \sup_{1 \leq N \leq R} |\mathcal{A}_{N;X,T} f(T^j x)|^p &= \sum_{j=0}^{J-R} \sup_{1 \leq N \leq R} |\mathcal{M}_N F(j)|^p \leq \sum_{j=0}^{J-R} \sup_{N \in \mathbb{N}} |\mathcal{M}_N F(j)|^p \\ &\leq \| \sup_{N \in \mathbb{N}} |\mathcal{M}_N f| \|_{\ell^p(\mathbb{Z})}^p \leq C_p^p \|F\|_{\ell^p(\mathbb{Z})}^p = C_p^p \sum_{j=0}^{J} |f(T^j x)|^p. \end{aligned}$$

Proof of the Calderón transference principle

Thus

$$\sum_{j=0}^{J-R} \int_X \sup_{1 \le N \le R} |\mathcal{A}_{N;X,T} f(T^j x)|^p d\mu(x) \le C_p^p \sum_{j=0}^{J} \int_X |f(T^j x)|^p d\mu(x).$$

Integrating both sides of this inequality, we get

$$\left(1 - \frac{R}{J}\right)^{1/p} \Big\| \sup_{1 \le N \le R} |\mathcal{A}_{N;X,T} f| \Big\|_{L^p(X)} \le C_p \|f\|_{L^p(X)},$$

taking $J \to \infty$ we obtain

$$\Big\| \sup_{1 \le N \le R} |\mathcal{A}_{N;X,T} f| \Big\|_{L^p(X)} \le C_p \|f\|_{L^p(X)},$$

Finally, taking $R \to \infty$ we have

$$\Big\| \sup_{N \in \mathbb{N}} \mathcal{A}_{N;X,T} f \Big\|_{L^p(X)} \le C_p \|f\|_{L^p(X)}.$$

The proof of the lemma is completed. □

Birkhoff's ergodic theorem

To establish that for every $1 \leq p < \infty$ and every $f \in L^p(X)$ there exists $f^* \in L^p(X)$ such that

$$\lim_{N \to \infty} A_{N;X,T} f(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) = f^*(x) \tag{6}$$

one can proceed in two steps:

- **Step 1.** Quantitative version of ergodic theorem

$$\| \sup_{N \in \mathbb{N}} |A_{N;X,T} f| \|_{L^p(X)} \lesssim \|f\|_{L^p(X)} \quad \text{for} \quad p \in (1, \infty]. \tag{7}$$

$$\mu(\{x \in X : \sup_{N \in \mathbb{N}} |A_{N;X,T} f(x)| > \lambda\}) \lesssim \lambda^{-1} \|f\|_{L^1(X)} \quad \text{for} \quad p = 1. \tag{8}$$

The bounds in (7) follow from the Hardy–Littlewood maximal inequality

$$\Big\| \sup_{N \in \mathbb{N}} \Big| \frac{1}{N} \sum_{n=1}^{N} f(x - n) \Big| \Big\|_{\ell^p(\mathbb{Z})} \lesssim \|f\|_{\ell^p(X)}, \quad \text{for} \quad p \in (1, \infty],$$

which is $A_{N,\mathbb{Z},S} f$ with $X = \mathbb{Z}$ and the shift operator $S(x) = x - 1$ in (6).

- **Step 2.** Pointwise convergence on a dense class of functions in $L^p(X)$.

Convergence on a dense class

$$A_N f(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^n x)$$

- $\mathbf{I}_T = \left\{ f \in L^2(X) : f \circ T = f \right\}$. If $f \in \mathbf{I}_T$, then

  $$A_N f = f,$$

  μ-almost everywhere.

- $\mathbf{J}_T = \left\{ g \circ T - g : g \in L^2(X) \cap L^\infty(X) \right\}$. If $f \in \mathbf{J}_T$, then by telescoping

  $$|A_N f(x)| = \frac{1}{N}\left| \sum_{n=1}^{N} g(T^{n+1}x) - g(T^n x) \right| = \frac{1}{N}|g(T^{N+1}x) - g(Tx)| \xrightarrow[N\to\infty]{} 0.$$

- $\mathbf{I}_T \oplus \mathbf{J}_T$ is dense in $L^2(X)$ by Riesz decomposition.

All together: pointwise convergence in $L^p(X)$

- Our aim will be to show that for any $f \in L^2(X)$ one has

$$\mu(\{x \in X : (A_{N;X,T}f(x))_{N\in\mathbb{N}} \text{ is not a Cauchy sequence}\}) = 0.$$

- Since $\mathbf{I}_T \oplus \mathbf{J}_T$ is dense in $L^2(X)$ we can find $(f_n)_{n\in\mathbb{N}} \subset \mathbf{I}_T \oplus \mathbf{J}_T$ so that $\|f_n - f\|_{L^2(X)} \xrightarrow[n\to\infty]{} 0$, and $\lim_{M,N\to\infty} |A_{M;X,T}f_n(x) - A_{N;X,T}f_n(x)| = 0$ for μ almost every $x \in X$.
- Suppose for a contradiction that there is $\delta > 0$ such that

$$\begin{aligned}\delta < \mu(\{x \in X : \limsup_{M,N\to\infty} |A_{M;X,T}f(x) - A_{N;X,T}f(x)| > \delta\}) \\ \leq \mu(\{x \in X : \sup_{N\in\mathbb{N}} |A_{N;X,T}(f_n - f)(x)| > \delta/2\}).\end{aligned}$$

- By the maximal inequality $\| \sup_{N\in\mathbb{N}} |A_{N;X,T}f| \|_{L^2(X)} \lesssim \|f\|_{L^2(X)}$ and Chebyshev inequality we obtain a contradiction, since one has

$$\delta < \mu(\{x \in X : \sup_{N\in\mathbb{N}} |A_{N;X,T}(f_n - f)(x)| > \delta/2\}) \leq \frac{C^2}{\delta^2} \|f_n - f\|_{L^2(X)}^2 \xrightarrow[n\to\infty]{} 0.$$

- For $p \neq 2$ we repeat the argument using the fact that $L^2(X) \cap L^p(X)$ is dense in $L^p(X)$ for any $1 \leq p < \infty$.

Bourgain's pointwise ergodic theorem

In the early 1980's Bellow and Furstenberg asked independently about the poinwise convergence for polynomial ergodic averages

$$A^P_{N;X,T}f(x) := \frac{1}{N}\sum_{n=1}^{N} f(T^{P(n)}x) \quad \text{for} \quad x \in X,$$

where $P \in \mathbb{Z}[\mathrm{n}]$ is a polynomial of degree > 1.

Bourgain used the circle method of Hardy and Littlewood to show:

- $L^p(X)$ boundedness of the maximal function for any $1 < p \le \infty$, i.e.

$$\|\sup_{N\in\mathbb{N}}|A^P_{N;X,T}f|\|_{L^p(X)} \lesssim \|f\|_{L^p(X)} \quad \text{for} \quad p \in (1,\infty].$$

- Given an increasing sequence $(N_j : j \in \mathbb{N})$, for each $J \in \mathbb{N}$ one has

$$\Big(\sum_{j=0}^{J}\big\|\sup_{N_j\le N<N_{j+1}}\big|A^P_{N;X,T}f - A^P_{N_j;X,T}f\big|\big\|^2_{L^2(X)}\Big)^{1/2} \le o(J^{1/2})\|f\|_{L^2(X)}.$$

Oscillation inequality for Birkhoff's operators

Recall that

$$A_{N;X,T}f(x) = \frac{1}{N}\sum_{n=1}^{N} f(T^n x).$$

Fix $\tau \in (1,2]$ and define $\Lambda = \{\lfloor \tau^k \rfloor : k \in \mathbb{N} \cup \{0\}\}$. Let $(k_j)_{j\in\mathbb{N}}$ be an increasing sequence and set $N_j = \lfloor \tau^{k_j} \rfloor$.

Theorem

Let $(X, \mathcal{B}(X), \mu, T)$ be a measure-preserving system then for every $J \in \mathbb{N}$ there is $C_J > 0$ such that we have

$$\sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big| A_{N;X,T}f - A_{N_j;X,T}f \big| \Big\|_{L^2(X)}^2 \leq C_J \|f\|_{L^2(X)}^2, \qquad (9)$$

and $\lim_{J\to\infty} C_J/J = 0$. In particular, for every $f \in L^2(X)$ there exists $f^* \in L^2(X)$ such that

$$\lim_{N\to\infty} A_{N;X,T}f(x) = f^*(x),$$

for μ-almost every $x \in X$.

Proof of the oscillation inequality for Birkhoff's operators

- Repeating the same argument as in the proof of transference principle it only suffices to work with $(\mathbb{Z}, \mathbf{P}(\mathbb{Z}), |\cdot|, S)$ with $S(x) = x - 1$. Then

$$A_{N;X,S}f(x) = M_N f(x) = \frac{1}{N}\sum_{n=1}^{N} f(x-n) = K_N * f(x), \qquad f \in \ell^2(\mathbb{Z}),$$

  where

$$K_N(x) = \frac{1}{N}\sum_{n=1}^{N} \delta_n(x), \qquad x \in \mathbb{Z}.$$

- By the bounds for the Hardy–Littlewood maximal function

$$\|\sup_{N\in\mathbb{N}} |M_N f|\|_{\ell^p(\mathbb{Z})} \lesssim \|f\|_{\ell^p(\mathbb{Z})}$$

  one can assume that $f \in \ell^2(\mathbb{Z}) \cap \ell^\infty(\mathbb{Z})$ and $f \geq 0$ is finitely supported.
- For $f \in \ell^1(\mathbb{Z})$ let us denote by

$$\hat{f}(\xi) = \sum_{n\in\mathbb{Z}} e^{2\pi i n\xi} f(n),$$

  the discrete Fourier transform on $\mathbb{Z}$ and let $\mathcal{F}^{-1}$ be its inverse.

Proof of the oscillation inequality for Birkhoff's operators

- One can see that $\widehat{M_N f}(\xi) = \hat{K}_N(\xi)\hat{f}(\xi)$, where

$$\hat{K}_N(\xi) = \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i n\xi}.$$

- Let $B_j = \{x \in (-1/2, 1/2) : |x| \le N_j^{-1}\}$. By Plancherel's theorem

$$\sum_{j=0}^{J} \Big\| \sup_{N\in\Lambda\cap(N_j,N_{j+1}]} \big|\mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j})\mathbb{1}_{B_{j+1}}\hat{f}\big)\big|\Big\|_{\ell^2}^2$$

$$\le \sum_{j=0}^{J} \sum_{N\in\Lambda\cap(N_j,N_{j+1}]} \big\|\mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j})\mathbb{1}_{B_{j+1}}\hat{f}\big)\big\|_{\ell^2}^2$$

$$\le \Big\| \sum_{j=0}^{J} \mathbb{1}_{B_{j+1}} \sum_{N\in\Lambda\cap(N_j,N_{j+1}]} |\hat{K}_N - \hat{K}_{N_j}|^2 \Big\|_{L^\infty} \|f\|_{\ell^2}^2.$$

Proof of the oscillation inequality for Birkhoff's operators

- For $N \in \Lambda \cap (N_j, N_{j+1}]$ we have

$$|\hat{K}_N(\xi) - \hat{K}_{N_j}(\xi)| \lesssim |\xi|N,$$

hence

$$\sum_{j=0}^{J} \mathbb{1}_{B_{j+1}}(\xi) \sum_{N \in \Lambda \cap (N_j, N_{j+1}]} |\hat{K}_N(\xi) - \hat{K}_{N_j}(\xi)|^2$$

$$\lesssim |\xi|^2 \sum_{j=0}^{J} \mathbb{1}_{B_{j+1}}(\xi) \sum_{N \in \Lambda \cap (N_j, N_{j+1}]} N^2$$

$$\lesssim |\xi|^2 \sum_{j : N_{j+1} \le |\xi|^{-1}} N_{j+1}^2 \lesssim 1.$$

Therefore, we obtain

$$\sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big| \mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j}) \mathbb{1}_{B_{j+1}} \hat{f} \big) \big| \Big\|_{\ell^2}^2 \lesssim \|f\|_{\ell^2}^2.$$

Proof of the oscillation inequality for Birkhoff's operators

- Similar for B_j^c, replacing $\hat{K}_{N_j}$ by $\hat{K}_{N_{j+1}}$ under the supremum, we can estimate

$$\sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big| \mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j}) \mathbb{1}_{B_j^c} \hat{f} \big) \big| \Big\|_{\ell^2}^2$$

$$\lesssim \sum_{j=0}^{J} \sum_{N \in \Lambda \cap [N_j, N_{j+1}]} \big\| \mathcal{F}^{-1}\big((\hat{K}_{N_{j+1}} - \hat{K}_N) \mathbb{1}_{B_j^c} \hat{f} \big) \big\|_{\ell^2}^2$$

$$\leq \Big\| \sum_{j=0}^{J} \mathbb{1}_{B_j^c} \sum_{N \in \Lambda \cap [N_j, N_{j+1}]} |\hat{K}_{N_{j+1}} - \hat{K}_N|^2 \Big\|_{L^\infty} \|f\|_{\ell^2}^2.$$

Now for $N \in \Lambda \cap [N_j, N_{j+1}]$ we obtain

$$|\hat{K}_{N_{j+1}}(\xi) - \hat{K}_N(\xi)| \lesssim |\xi|^{-1} N^{-1}$$

Proof of the oscillation inequality for Birkhoff's operators

- Thus

$$\sum_{j=0}^{J} \mathbb{1}_{B_j^c}(\xi) \sum_{N\in\Lambda\cap[N_j,N_{j+1}]} |\hat{K}_{N_{j+1}}(\xi) - \hat{K}_N(\xi)|^2$$

$$\lesssim |\xi|^{-2} \sum_{j=0}^{J} \mathbb{1}_{B_j^c}(\xi) \sum_{N\in\Lambda\cap[N_j,N_{j+1}]} N^{-2}$$

$$\lesssim |\xi|^{-2} \sum_{j:N_j\geq|\xi|^{-1}} N_j^{-2} \lesssim 1.$$

Therefore, we conclude

$$\sum_{j=0}^{J} \Big\| \sup_{N\in\Lambda\cap(N_j,N_{j+1}]} \big|\mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j})\mathbb{1}_{B_j^c}\hat{f}\big)\big|\Big\|_{\ell^2}^2 \lesssim \|f\|_{\ell^2}^2.$$

Proof of the oscillation inequality for Birkhoff's operators

- Finally, for $p = 2$ we obtain

$$\sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big| \mathcal{F}^{-1}\big((\hat{K}_N - \hat{K}_{N_j}) \mathbb{1}_{B_j} \mathbb{1}_{B_{j+1}^c} \hat{f} \big) \big| \Big\|_{\ell^2}^2$$

$$\lesssim \sum_{j=0}^{J} \big\| \mathcal{F}^{-1}\big(\mathbb{1}_{B_j} \mathbb{1}_{B_{j+1}^c} \hat{f} \big) \big\|_{\ell^2}^2 \leq \|f\|_{\ell^2}^2.$$

- Hence

$$\sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big| A_{N;X,T} f - A_{N_j;X,T} f \big| \Big\|_{L^2(X)}^2 \leq C_J \|f\|_{L^2(X)}^2,$$

- In fact we have proved that C_J is constant. The proof of (9) is completed.

How oscillations imply pointwise convergence

- By the maximal inequality for $p = 2$ we can assume that $f \in L^2(X)$ is bounded and $\|f\|_{L^\infty(X)} \leq 1$.
- Suppose for a contradiction that $(A_{N;X,T}f(x))_{N\in\mathbb{N}}$ does not converge. Then there is $\varepsilon \in (0,1)$ such that

$$\mu\Big(\{x \in X : \limsup_{M,N\to\infty} |A_{M;X,T}f(x) - A_{N;X,T}f(x)| > 8\varepsilon\}\Big) > 8\varepsilon.$$

- Thus there exists $(k_j)_{j\in\mathbb{N}}$ such that

$$\mu\Big(\{x \in X : \sup_{N_j<N\leq N_{j+1}} |A_{N;X,T}f(x) - A_{N_j;X,T}f(x)| > 4\varepsilon\}\Big) > 4\varepsilon,$$

  where $N_j = \lfloor \tau^{k_j} \rfloor$ and $\tau = 1 + \varepsilon/4$.
- If $\lfloor \tau^k \rfloor \leq N < \lfloor \tau^{k+1} \rfloor$ then

$$|A_{N;X,T}f - A_{\lfloor \tau^k \rfloor;X,T}f| = \left| \frac{1}{N} \sum_{n=\lfloor \tau^k \rfloor+1}^{N} f(T^n x) - \frac{N - \lfloor \tau^k \rfloor}{N\lfloor \tau^k \rfloor} \sum_{n=1}^{\lfloor \tau^k \rfloor} f(T^n x) \right|$$

$$\leq \frac{2(N - \lfloor \tau^k \rfloor)}{N} \leq \frac{4\tau^k(\tau-1)}{\tau^k} + \frac{4}{\tau^k} = 4(\tau - 1) + \frac{4}{\tau^k} < 2\varepsilon,$$

  for $k \geq k_0$, since we can always arrange k_0 to satisfy $\tau^{-k_0} < \varepsilon/4$.

How oscillations imply pointwise convergence

- Therefore, we obtain that

$$\mu\Big(\{x \in X : \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} |A_{N;X,T}f(x) - A_{N_j;X,T}f(x)| > \varepsilon\}\Big) > \varepsilon.$$

- Now applying oscillation inequality we obtain that

$$0 < \varepsilon^3 \leq \frac{1}{J} \sum_{j=0}^{J} \Big\| \sup_{N \in \Lambda \cap (N_j, N_{j+1}]} \big|A_{N;X,T}f - A_{N_j;X,T}f\big| \Big\|_{L^2(X)}^2 \leq J^{-1} C_J \|f\|_{L^2(X)}^2,$$

  but it is impossible since, the right-hand side tends to 0 as $J \to \infty$.
- This proves the pointwise convergence of $A_{N;X,T}f$ on $L^2(X)$ and completes the proof.

Dziękuję!