ILLINOIS JOURNAL OF MATHEMATICS
Volume 44, Number 4, Winter 2000

A SMOOTHER ERGODIC AVERAGE

KARIN REINHOLD

ABSTRACT. We study the pointwise behavior of the smoothed out averages

$$P_n f(x) = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{\epsilon_k} \int_{|t|<\epsilon_k/2} f(T_{k+t}x)dt,$$

where T_t is a measure preserving flow on a probability space. We show that these are good averages in L^p, $p > 1$, if ϵ_k is a convergent sequence or if they are given by stationary random variables. When $p = 1$ the averages are good if $\lim_{k\to\infty} \epsilon_k = \epsilon > 0$.

1. Introduction

Let (X, β, m) be a probability space and T a measure preserving point transformation on it. Birkhoff's Ergodic Theorem shows that the averages $A_n f(x) = (1/n) \sum_{k=1}^{n} f(T^k x)$ converge almost everywhere for any function $f \in L^1(X)$. However Bergelson, Boshernitzan and Bourgain [3] showed that if $\{T_t\}$ is a flow on X given by measure preserving point transformations, then the averages along observations slightly perturbed at time k, $B_n f(x) = (1/n) \sum_{k=1}^{n} f(T_{k+\epsilon_k}x)$, fail to converge even for $L^\infty(X)$ functions. In fact these averages fail to converge so badly that Akcoglu, Bellow, del Junco and Jones [1] showed that they are strong sweeping out averages for any sequence $\epsilon_k \to 0$.

They concluded then that in applications, the ergodic averages do not converge because, by limitations of instruments, time measurements can not be taken exactly at instant k. Therefore, instead of dealing with averages along arithmetic sequences (as is the case of the averages $A_n f$), in applications one rather has an average of the form $B_n f$. However, in the same spirit, one could also argue that by limitation of instruments one can not measure time instantly. That is, measurements can not be taken at any particular instant, rather they are an average measurement around those instants. One such model results when the measurement at time k is actually the following average measurement around instant k:

$$\frac{1}{\epsilon_k} \int_{|t|<\epsilon_k/2} f(T_{k+t}x)dt.$$

Received May 14, 1999; received in final form November 18, 1999.
1991 Mathematics Subject Classification. Primary 47A35, 28Dxx.
The author was supported by a grant from the National Science Foundation.

In this model, the time averages take the form

$$P_n f(x) = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\epsilon_k}\int_{|t|<\epsilon_k/2} f(T_{k+t}x)dt. \tag{1}$$

These averages have extra smoothness and therefore are expected to be better behaved. And indeed they are.

In this work we look at a more general setting than in (1). That is, we assume that the averages at time k are "weighted" averages, where the weights are given by an approximation to the identity.

The paper is organized as follows. In Section 2 we show that maximal function associated with these type of averages is strong type (p,p) for $p > 1$, and when $\inf_k \epsilon_k > 0$, the maximal function is also weak type (1,1). We also show that the averages converge if the sequence ϵ_k converge.

The convergence of the sequence ϵ_k is not necessary for the convergence of these averages. In Section 3 we study a certain kind of non–convergent sequences ϵ_k for which there is pointwise convergence of $P_n f(x)$. However, in these cases, the ergodicity of T_1 is necessary. A counterexample also shows that in the general case, ergodicity of T_1 is not sufficient either.

Lastly, Section 4 shows convergence results when the sequence ϵ_k is given by a stationary sequence of random variables.

We thank the referee for very useful comments.

Throughout this work, a probability space means a complete probability space. And C denotes a constant whose value may change from one occurrence to the next.

2. Positive results

Let $\{T_t\}_{t\in\mathbb{R}}$ be a measurable flow of measure preserving transformation on a probability space (X, β, m). Let φ be a positive, integrable function on $\mathbb{R}$ satisfying the following conditions:

(a) $\int \varphi(x)dx = 1$.
(b) The function $\phi(x) = \sup_{|y|\geq|x|}\varphi(y)$ is integrable.

Define $\varphi_\epsilon f(x) = \int f(T_t x)\varphi_\epsilon(t)dt$, where $\varphi_\epsilon(t) = \varphi(t/\epsilon)/\epsilon$. Given any sequence of positive numbers $\{\epsilon_k\}_{k=1}^{\infty} \subset (0, 1)$, define the averages

$$P_n f(x) = \frac{1}{n}\sum_{k=1}^{n}\varphi_{\epsilon_k} f(T_k x).$$

We will also consider the maximal operator $Mf(x) = \sup_{n\geq1}|P_n f(x)|$.

LEMMA 2.1. *Consider $\mathbb{R}$ with Lebesgue measure and the shift flow $T_t x = x + t$. Then there exists a universal constant C such that:*

(A) *For any $f \in L^p(\mathbb{R})$, $p > 1$,*

$$\|Mf\|_p \leq \frac{C}{(p-1)^2}\|f\|_p,$$

and for any $f \in L \log L$,

$$m\{x\colon Mf(x) > \lambda\} \leq \frac{C}{\lambda}\int |f|\left(1 + \log^+ \frac{|f|}{\|f\|_1}\right) dm.$$

(B) *If $\inf_{k\geq 1} \epsilon_k \geq \epsilon > 0$, then for any $f \in L^1(\mathbb{R})$,*

$$m(\{x\colon Mf(x) > \lambda\}) \leq \frac{C}{\epsilon\lambda}\|f\|_1.$$

Proof. This lemma is immediate by the Maximal Ergodic Theorem and the Hardy–Littlewood Maximal Theorem because under our assumptions on φ, the maximal function $\sup_{\delta>0}\varphi_\delta f$ is dominated by the Hardy–Littlewood maximal function $\sup_{\epsilon>0}(1/\epsilon)\int_{|t|<\epsilon/2} f(x+t)dt$ (see [15] or [16]). Thus $\|\sup_n |A_n f|\|_p \leq (C/(p-1))\|f\|_p$ and $\|\sup_{\delta>0}\varphi_\delta f\|_p \leq (C/(p-1))\|f\|_p$ for $p > 1$, where C denotes a universal constant.

(A) Since $|P_n f| \leq A_n(\sup_{\delta>0}|\varphi_\delta f|)$, then, for $f \in L^p$ with $p > 1$,

$$\left\|\sup_n |P_n f|\right\|_p \leq \frac{C}{(p-1)}\|\sup_{\delta>0}\varphi_\delta f\|_p \leq \frac{C'}{(p-1)^2}\|f\|_p.$$

Now, if $f \in L\log L$, then $\sup_{\delta>0}|\varphi_\delta f| \in L^1$. Hence

$$\begin{aligned} m\left\{s\colon \sup_n |P_n f(s)| > \lambda\right\} &\leq m\left\{s\colon \sup_n |A_n(\sup_{\delta>0}|\varphi_\delta f(s)|) > \lambda\right\} \\ &\leq \frac{C}{\lambda}\left\|\sup_{\delta>0}|\varphi_\delta f|\right\|_1 \leq \frac{C}{\lambda}\int |f|\left(1 + \log^+\frac{|f|}{\|f\|_1}\right) dm. \end{aligned}$$

(B) Since $\epsilon < \epsilon_k < 1$,

$$|\varphi_{\epsilon_k} f(x)| \leq \frac{1}{\epsilon}\int |f|(x+t)\varphi(t/\epsilon_k)dt \leq \frac{1}{\epsilon}\int |f|(x+t)\phi(t)dt.$$

Letting $Sf(x) = \int f(x+t)\phi(t)dt$, we have

$$\begin{aligned} m\left(\left\{x\colon \sup_n |P_n f(x)| > \lambda\right\}\right) &\leq m\left(\left\{x\colon \sup_n A_n S|f|(x) > \epsilon\lambda\right\}\right) \\ &\leq \frac{C}{\epsilon\lambda}\|S|f|\|_1 = \frac{C'}{\epsilon\lambda}\|f\|_1. \end{aligned}$$

□

The next lemma is essentially Calderón Transfer Lemma [6], adapted to our needs.

LEMMA 2.2. (A) *If Mg is a weak* (1,1) *operator for functions g on $\mathbb{R}$ with Lebesgue measure and flow $T_t x = x + t$, then for any measure preserving flow $\{T_t\}$ on a probability space (X, β, m), the operator Mf is also weak* (1,1).

(B) *If Mg is a strong (p, p), $1 < p \leq \infty$, operator for functions g on $\mathbb{R}$ with Lebesgue measure and flow $T_t x = x + t$, then for any measure preserving flow $\{T_t\}$ on a probability space (X, β, m), the operator Mf is also strong (p, p).*

LEMMA 2.3. *Let $\{\epsilon_k\}$ be a sequence of real numbers in* (0, 1).

(a) *Then $\sup_{n\geq 1} P_n f$ satisfies a strong (p,p) inequality.*
(b) *If $\epsilon_k \geq \epsilon > 0$ for all k, then $\sup_{n\geq 1} P_n f$ satisfies a weak* (1,1) *inequality.*

Proof. This lemma follows from Lemmas 2.1 and 2.2. □

PROPOSITION 2.4. *Given $f \in L^1(X)$, then $\lim_{\epsilon\to 0} \varphi_\epsilon f(x) = f(x)$ for a.e. x.*

Proof. Let $B = \{x\colon \varphi_\epsilon f(x) \to f(x) \text{ as } \epsilon \to 0\}$.

Since $\{T_t\}$ is a measure preserving, measurable flow, the application $\Psi\colon X\times\mathbb{R} \longrightarrow X$, $\Psi(x, t) = T_t x$ is measurable and $m(T_t A) = m(A)$ for all $A \in \beta$. With this notation, let

$$\Psi^{-1}(B) = \{(x, t)\colon \varphi_\epsilon f(T_t x) \to f(T_t x) \text{ as } \epsilon \to 0\}$$

Let $C = \Psi^{-1}(B^c)$, the set of pairs (x, t) where convergence fails.

Let $C^x = \{t\colon (x, t) \in C\}$ and $C_t = \{x\colon (x, t) \in C\}$ and let λ denote the Lebesgue measure on $\mathbb{R}$. Then, by Fubini's Theorem, $\lambda(C^x) = 0$ for m—a.e. x.

Indeed, consider $F(x, t) = f(\Psi(x, t)) = f(T_t x)$. Note that $\varphi_\epsilon F_x(t) = \int F_x(t + s)\varphi_\epsilon(s)ds$. Now, since f is integrable and Ψ is measurable, F is measurable and locally integrable. By Fubini's Theorem, $F_x(t) = F(x, t)$ is measurable and locally integrable for m—a.e. x. For such an x, the Lebesgue Differentiation Theorem adapted to convolutions with approximation to the identity (see [15] or [16]), gives $\lim_{\epsilon\to 0} \varphi_\epsilon F_x(t) = F_x(t)$ for a.e. t, proving the claim.

Then

$$0 = \int_X \lambda(C^x)dm(x) = m \times \lambda(C) = \int_R m(C_t)dt.$$

Since $0 \leq m(C_t) \leq 1$ for a.e. t, we have $m(C_t) = 0$ for a.e. t. However

$$C_t = \{x\colon \Psi(x, t) \in B^c\} = \{x\colon T_t x \in B^c\} = T_t^{-1}(B^c).$$

Since the flow is measure preserving, $m(B^c) = 0$. Hence $\lim_{\epsilon\to 0} \varphi_\epsilon f(x) = f(x)$ for a.e. x. □

Let $f_\epsilon = \varphi_\epsilon f$ if $\epsilon > 0$ and $f_0 = f$. Proposition 2.4 has the following immediate consequence.

COROLLARY 2.5. *If* $\lim_{k\to\infty} \epsilon_k = \epsilon \geq 0$, *then* $\lim_{n\to\infty}(1/n)\sum_{k=1}^{n} \varphi_{\epsilon_k} f(x) = f_\epsilon(x)$ *a.e. for any* $f \in L^1$.

COROLLARY 2.6. *If* $1 \leq p < \infty$, *then* $\lim_{\epsilon\to 0} \varphi_\epsilon f = f$ *in* L^p.

Proof. Let $f \in L^p$. Given $\eta > 0$ there exists $\delta > 0$ such that if $m(A) < \delta$ then $\int_A |f|^p dm < \eta/2^p$. This property also shows that for any ϵ, $\int_A |\varphi_\epsilon f|^p dm < \eta/2^p$. Indeed, since the flow is measure preserving, $m(T_t^{-1}A) = m(A) < \delta$ for all t and hence

$$\begin{aligned}
\int_A |\varphi_\epsilon f(x)|^p dm(x) &\leq \int_A \int |f(T_t x)|^p \varphi_\epsilon(t) dt dm(x) \\
&= \int \varphi_\epsilon(t) \int_A |f(T_t x)|^p dm(x) dt \\
&= \int \varphi_\epsilon(t) \int_{T_t^{-1}A} |f(x)|^p dm(x) dt \\
&< \frac{\eta}{2^p} \int \varphi_\epsilon(t) dt = \frac{\eta}{2^p}.
\end{aligned}$$

Let $A_{\eta,\nu} = \{x\colon\ |\varphi_\epsilon f(x) - f(x)| < \eta \text{ for all } \epsilon < \nu\}$. By Proposition 2.4, $m(A_{\eta,\nu}) \to 1$ as $\nu \to 0$. Fix $\nu = \nu(\eta)$ so that $m(A_{\eta,\nu}) > 1 - \delta$. Then, given $\eta > 0$ there exists ν such that for all $\epsilon < \nu$,

$$\begin{aligned}
\|\varphi_\epsilon f - f\|_p^p &= \int_{A_{\eta,\nu}} |\varphi_\epsilon f - f|^p dm + \int_{A_{\eta,\nu}^c} |\varphi_\epsilon f - f|^p dm \\
&\leq \eta + 2^p \int_{A_{\eta,\nu}^c} |\varphi_\epsilon f|^p dm + 2^p \int_{A_{\eta,\nu}^c} |f|^p dm \\
&\leq 3\eta.
\end{aligned}$$

□

To prove the pointwise convergence of the averages $P_n f$, we refer to the following theorem of R. Jones and M. Wierdl (Theorem 2.10 in [9]).

THEOREM 2.7. *Let* (X, Σ, m) *be a measure space. Denote by* D *the set of a.e. measurable functions. Let* φ *be a Young function and let* $(L_\varphi, \|\ \|_\varphi)$ *be the corresponding Orlicz space. Let* $T_{n,k}$ *be a dissipative double sequence of linear operators which are continuous in measure, positive, and map* $L_\varphi \to D$. *Let* $\tau_n = \sum_k T_{n,k}$ *be the associated averaging sequence. Assume for every* $f \in L_\varphi$ *we have* $\tau_n f \to \tilde{f}$ *a.e. for some* $\tilde{f} \in D$. *Suppose that* $(g_k) \subset L_\varphi$ *is a sequence of functions with the* *property that* $g_k \to g$ *a.e. for some* $g \in L_\varphi$, *and* $\| \sup_k |g_k| \|_\varphi < \infty$; *then*

$$\sum_{k=1}^{\infty} T_{n,k} g_k \to \tilde{g} \quad a.e.$$

Let πf denote the projection of f onto the subspace of invariant functions under T_1 defined by $\lim_{n\to\infty} A_n(f)(x) = \pi f(x)$ for a.e. x and $f \in L^p$, $p \geq 1$. We are now ready to prove the convergence of the averages $P_n f$.

PROPOSITION 2.8. *Let* $\{T_t\}$ *be a measure preserving flow on a probability space* (X, β, m).

(a) *If* $\lim_{k\to\infty} \epsilon_k = \epsilon > 0$, *then for any* $f \in L^p$, $p \geq 1$, $\lim_{n\to\infty} P_n f(x) = \pi f_\epsilon$ *for a.e.* x.
(b) *If* $\lim_{k\to\infty} \epsilon_k = 0$ *then for any* $f \in L^p$, $p > 1$, $\lim_{n\to\infty} P_n f(x) = \pi f$ *a.e.* x.

Proof. This proposition is a consequence of Jones and Wierdl's Theorem 2.7. In our situation, $L_\varphi = L^p$, $T_{n,k} f(x) = f(T_k x)/n$ and $\tau_n f(x) = A_n f(x) = n^{-1} \sum_{k=1}^{n} f(T_k x)$. Clearly, $T_{n,k}$ is dissipative since, for any $f \in L^p$, $\lim_{n\to\infty} f(T_k x)/n = 0$ for a.e. x.

By the Pointwise Ergodic Theorem $\tau_n f \to \pi f$ a.e. for any $f \in L^p$. The functions $g_k = \varphi_{\epsilon_k} f$ are all in L^p if $f \in L^p$, and by Proposition 2.4, have the property $g_k \to f_\epsilon$ a.e..

It remains only to show that $\sup_k |g_k| \in L^p$.

We have two cases to consider. If $\lim_{k\to\infty} \epsilon_k = \epsilon > 0$, then $\epsilon_k \geq \epsilon_0 = \inf_k \epsilon_k > 0$. Hence

$$\sup_k |g_k| \leq \sup_{\delta \geq \epsilon_0} \varphi_\delta f \leq \frac{1}{\epsilon_0} \int f(T_t x) \phi(t) dt.$$

Since ϕ is integrable, $\sup_k |g_k| \in L^p$ if $f \in L^p$, $p \geq 1$.

If $\lim_{k\to\infty} \epsilon_k = 0$, then (see [15] or [16]) the maximal function is dominated by the Hardy–Littlewood maximal function

$$\sup_k |g_k| \leq \sup_{\epsilon > 0} \varphi_\epsilon |f| \leq C(\varphi) \sup_{\epsilon > 0} \frac{1}{\epsilon} \int_{|t| < \epsilon} |f(x+t)| dt,$$

and the later maximal function is in L^p, only if $p > 1$.

In either case, Theorem 2.7 gives $P_n f = n^{-1} \sum_{k=1}^{n} T_k g_k \to \pi f_\epsilon$ a.e.. □

The technique used in Proposition 2.8 failed to show pointwise convergence in L^1 in the case when $\epsilon_k \to 0$. However, mean convergence holds even in this case. It is not difficult to see that if $\lim_{k\to\infty} \epsilon_k = \epsilon \geq 0$, there is mean convergence in L^p for $p \geq 1$. Indeed, mean convergence in L^p for $p > 1$ or in L^1 when $\epsilon_k \to \epsilon > 0$, is an immediate consequence of Proposition 2.8. The next corollary then completes the case $p = 1$, showing that if $\epsilon_k \to 0$ then the averages $P_n f$ converge at least in the L^1 norm.

COROLLARY 2.9. *Let* $1 \leq p < \infty$. *If* $\lim_{k\to\infty} \epsilon_k = \epsilon \geq 0$, *then* $\lim_{n\to\infty} P_n f$ *converges to* πf_ϵ *in* L^p.

This corollary is an application of the following result and Corollary 2.6.

LEMMA 2.10. *Let* $\{T_{n,k}\}$ *be a double sequence of operators on a Banach space* $(Y, \|.\|)$ *such that*

(a) $\|T_{n,k}\| \to 0$ *as* $n \to \infty$,
(b) $\sup_n \sum_{k=1}^{\infty} \|T_{n,k}\| = C < \infty$,

where $\|T_{n,k}\|$ *denotes the operator norm. Let* $\tau_n = \sum_{k=1}^{\infty} T_{n,k}$ *and assume that* $\lim_{n\to\infty} \tau_n y = \bar{y}$ *for any* $y \in Y$. *Let* $y_k, y \in Y$ *such that* $\lim_{k\to\infty} y_k = y$. *Then* $\lim_{n\to\infty} \sum_{k=1}^{\infty} T_{n,k} y_k = \bar{y}$.

Proof. Since $\tau_n y \to \bar{y}$, it suffices to show that $\sum_{k=1}^{\infty}(T_{n,k}y_k - T_{n,k}y) \to 0$ as $n \to \infty$.

$$\begin{aligned}\left\|\sum_{k=1}^{\infty} T_{n,k}y_k - \sum_{k=1}^{\infty} T_{n,k}y\right\| &\leq \sum_{k=1}^{\infty} \|T_{n,k}(y_k - y)\| \leq \sum_{k=1}^{\infty} \|T_{n,k}\|\|y_k - y\| \\ &\leq \sum_{k=1}^{K} \|T_{n,k}\|\|y_k - y\| + \sum_{k=K+1}^{\infty} \|T_{n,k}\|\|y_k - y\| \\ &= I + II.\end{aligned}$$

To handle the second term, choose K so that given $\epsilon > 0$, $\|y_k - y\| < \epsilon/C$ for all $k > K$. Hence $II < (\epsilon/C)\sum_{k=K+1}^{\infty} \|T_{n,k}\| \leq (\epsilon/C)\, C = \epsilon$.

For that fixed K, since $\|T_{n,k}\| \to 0$ as $n \to \infty$, given $\epsilon > 0$ there exists $N \geq K$ such that $I < \epsilon$ for all $n \geq N$.

Thus we have shown that for any $\epsilon > 0$ there exist N such that $\| \sum_{k=1}^{\infty} T_{n,k}y_k - \sum_{k=1}^{\infty} T_{n,k}y\| \leq 2\epsilon$ for all $n \geq N$. □

3. Non-convergent sequences

In this section we only consider the special case $\varphi = \chi_{[-.5,.5]}$. To avoid confusion we will denote $\varphi_\epsilon f$ by $I_\epsilon f$. We will also use the notation $P_n' f(x) = (1/n)\sum_{k=1}^{n} I_{\epsilon_k} f(x)$.

LEMMA 3.1. *Let* $D = \operatorname{span}\{f - T_1 f\colon f \in L^\infty\}$. *If* $\lim_{n\to\infty}(1/n)\sum_{k=2}^{n} |\epsilon_{k-1} - \epsilon_k|/(\epsilon_{k-1} \vee \epsilon_k) = 0$, *then* $\lim_{n\to\infty} P_n f$ *exists and* $= 0$ *a.e. for any* $f \in D$.

Proof. Let $f = g - T_1 g$, $g \in L^\infty$. Then

$$P_n f(x) = \frac{1}{n}\sum_{k=2}^{n} I_{\epsilon_k} g(T_k x) - I_{\epsilon_{k-1}} g(T_k x) + \frac{I_{\epsilon_1} g(T_1 x)}{n} - \frac{I_{\epsilon_n} g(T_{n+1} x)}{n}.$$

Assume $\epsilon_{k-1} \geq \epsilon_k$; then

$$\begin{aligned}
&|I_{\epsilon_k} g(x) - I_{\epsilon_{k-1}} g(x)| \\
&= \left| \left(\frac{1}{\epsilon_k} - \frac{1}{\epsilon_{k-1}} \right) \epsilon_k I_{\epsilon_k} g(x) + \frac{1}{\epsilon_{k-1}} \int_{\epsilon_k \leq |t| < \epsilon_{k-1}} g(T_t x) dt \right| \\
&\leq 2 \frac{(\epsilon_{k-1} - \epsilon_k)}{\epsilon_{k-1}} \|g\|_\infty .
\end{aligned}$$

Thus

$$|P_n f(x)| \leq \frac{2}{n} \sum_{k=2}^{n} \frac{|\epsilon_{k-1} - \epsilon_k|}{\epsilon_{k-1} \vee \epsilon_k} \|g\|_\infty + \frac{2\|g\|_\infty}{n} \longrightarrow 0$$

as $n \to \infty$. $\square$

LEMMA 3.2. *Let $\{\epsilon_k\}$ be a sequence in* (0,1) *satisfying the variation condition* $\lim_{n\to\infty}(1/n)\sum_{k=2}^{n} |\epsilon_{k-1} - \epsilon_k|/(\epsilon_{k-1} \vee \epsilon_k) = 0$. *Then* $\lim_{n\to\infty} P'_n f(x)$ *exists a.e. for any $f \in L^p$ invariant under T_1 if and only if the following hold*:

(a) *If $p > 1$, $\lim_{n\to\infty} P_n f$ exists a.e. for any $f \in L^p$*,
(b) *If $p = 1$ and $\epsilon_k \geq \epsilon > 0$ for all k, $\lim_{n\to\infty} P_n f$ exists a.e. for any $f \in L^1$*.

Proof. Let $I = \{f \in L^p \colon T_1 f = f\}$, the invariant functions of T_1, and $D = \operatorname{span}\{f - T_1 f \colon f \in L^\infty\}$. The direction ($\Leftarrow$) is trivial because if $f \in I$ then $P'_n f(x) = P_n f(x)$. Let's prove ($\Rightarrow$). By Lemma 3.2, if $f \in D$, $\lim_{n\to\infty} P_n f = 0$ a.e.. On the other hand if $f \in I$, then $P_n f(x) = P'_n f(x)$ which converges a.e. by hypothesis. Hence $P_n f(x)$ converges a.e. for any function in $I + D$, a dense subspace of L^p. The lemma then follows by the maximal inequality in Lemma 2.3. $\square$

Lemma 3.2 is interesting because it says that if the sequence $\{\epsilon_k\}$ doesn't have a large variation, then $P_n f$ converges for all f in L^p provided $P' f$ converges for invariant functions. In the case when T_1 is ergodic, the second statement is trivial since the only invariant functions are the constants. But if the sequence $\{\epsilon_k\}$ does not converge, then ergodicity of T_1 is necessary.

LEMMA 3.3. *T_1 is ergodic if and only if for any sequence $\{\epsilon_k\}$ satisfying the condition of Lemma* 3.2 *one has the following*:

(a) *If $p > 1$, $\lim_{n\to\infty} P_n f$ exists a.e. for any $f \in L^p$*.
(b) *If $p = 1$ and $\epsilon_k \geq \epsilon > 0$ for all k, $\lim_{n\to\infty} P_n f$ exists a.e. for any $f \in L^1$*.

Proof. ($\Rightarrow$) The proof follows from Lemma 3.2 because if T_1 is ergodic then the only invariant functions are the constants so $P'_n f(x)$ converges trivially for any invariant function.

($\Leftarrow$) If T_1 is not ergodic, we will construct a sequence for which the averages do not converge. Construct a sequence $\{\epsilon_k\}$ taking only two values α and β in the following way. Let $a_1 = 1$ and inductively construct a_k and b_k such that $b_k = 2a_k$ and $a_{k+1} = 2b_k$. Then the collection of sets $A_k = [a_k, b_k)$ and $B_k = [b_k, a_{k+1})$ form a partition of the positive integers. Now let $\epsilon_j = \alpha$ if $j \in A_k$ and $\epsilon_j = \beta$ if $j \in B_k$.

Then, if f is an invariant function for T_1,

$$\begin{aligned} P_{b_k-1}f(x) &= I_\alpha\left(\frac{1}{b_k-1}\sum_{k\in\bigcup_{j\le k}A_j} f(T_1^n x)\right) + I_\beta\left(\frac{1}{b_k-1}\sum_{k\in\bigcup_{j<k}B_j} f(T_1^n x)\right) \\ &= \frac{\sum_{j\le k}|A_j|}{b_k-1} I_\alpha f(x) + \frac{\sum_{j<k}|B_j|}{b_k-1} I_\beta f(x). \end{aligned}$$

But

$$\begin{aligned} |A_k| &= b_k - a_k = a_k = 4^{k-1}, \\ |B_k| &= a_{k+1} - b_k = b_k = 24^{k-1}, \end{aligned}$$

so

$$\begin{aligned} \frac{\sum_{j\le k}|A_j|}{b_k-1} &= \frac{\sum_{j=0}^{k-1}4^j}{2\,4^{k-1}-1} = \frac{4^k-1}{3(2\,4^{k-1}-1)} \to \frac{2}{3} \text{ as } k\to\infty, \\ \frac{\sum_{j<k}|B_j|}{b_k-1} &= \frac{\sum_{j=0}^{k-2}2\,4^j}{2\,4^{k-1}-1} = \frac{4^{k-1}-1}{3(4^{k-1}-1/2)} \to \frac{1}{3} \text{ as } k\to\infty. \end{aligned}$$

Hence,

$$\lim_{k\to\infty} P_{b_k-1}f(x) = \frac{2}{3}I_\alpha f(x) + \frac{1}{3}I_\beta f(x).$$

Similarly,

$$\lim_{k\to\infty} P_{a_k-1}f(x) = \frac{1}{3}I_\alpha f(x) + \frac{2}{3}I_\beta f(x).$$

By choosing an invariant function f which is not constant, we see that the limit of $P_n f$ does not exists a. e.

Clearly the sequence ϵ_k satisfies the conditions of Lemma 3.2 because in a block of length 2^n there are only n changes. Thus this example shows that ergodicity is necessary. □

Example 3.4. With the help of Lemma 3.3 we can construct non–convergent sequences $\{\epsilon_k\}$ (taking more than finitely many values as in the above example) for which the averages P_n converge a.e. when T_1 is ergodic.

Construction. Fix $0 < \epsilon < 1/4$ and find $n > 1/\epsilon$. Construct the sequence ϵ_k inductively. Let $\epsilon_1 = \epsilon$ and

$$\epsilon_{k+1} = \begin{cases} \min\{\epsilon_k + \frac{1}{n+k}, \frac{1}{2}\} & \text{if } \epsilon_{k-1} < \epsilon_k < 1/2 \text{ or } \epsilon_k = \epsilon, \\ \max\{\epsilon, \epsilon_k - \frac{1}{n+k}\} & \text{if } \epsilon < \epsilon_k < \epsilon_{k-1} \text{ or } \epsilon_k = 1/2. \end{cases}$$

This sequence satisfies the hypothesis of Lemma 3.3, thus $P_n f(x)$ converges a.e. for $f \in L^1$, but the sequence $\{\epsilon_k\}$ itself does not converge.

The results obtained so far in Sections 2 and 3 seem to indicate that if T_1 is ergodic, given any sequence $\{\epsilon_k\}$ in (0,1), the averages $P_n f(x)$ converges a.e. for any $f \in L^p$, $p > 1$. However the ergodicity of T_1 alone is not sufficient. We thank the referee for the following example.

Example 3.5. Consider the flow $T_t x = x + \sqrt{2}t$ on the Torus [0,1] mod 1. Let $f(x) = 1$ if $x \in [0, 1/2)$ and $f(x) = -1$ if $x \in [1/2, 1)$. Let $\{x\}$ denote the fractional part of x.

Let $\{N_j\}_{j=0}^{\infty}$ be a rapidly growing sequence of integers to be determined, say $N_0 = 1$ and $N_{j+1} \geq 10N_j$ for $j \geq 1$. Let $\epsilon = \sqrt{2}/200$ and $J_1 = (\epsilon, \frac{1}{2} - 2\epsilon)$ and $J_2 = (\frac{1}{2} + \epsilon, 1 - 2\epsilon)$.

Define the sequence $\{\epsilon_k\}_{k=1}^{\infty}$ in the following fashion:

If $N_{2j} \leq k < N_{2j+1}$, let

$$\epsilon_k = \begin{cases} \frac{1}{100} & \text{if } \{k\sqrt{2}\} \in J_2 \\ \frac{1}{\sqrt{2}} & \text{if } \{k\sqrt{2}\} \notin J_2. \end{cases}$$

If $N_{2j+1} \leq k < N_{2j+2}$, let

$$\epsilon_k = \begin{cases} \frac{1}{100} & \text{if } \{k\sqrt{2}\} \in J_1 \\ \frac{1}{\sqrt{2}} & \text{if } \{k\sqrt{2}\} \notin J_1. \end{cases}$$

Notice that for all x,

$$I_{1/\sqrt{2}} f(x) = \sqrt{2} \int_{-1/2\sqrt{2}}^{1/2\sqrt{2}} f(x + \sqrt{2}t)dt = \int_{-1/2}^{1/2} f(x+t)dt = 0.$$

And

$$\begin{aligned} I_{1/100} f(x) &= 100 \int_{-1/200}^{1/200} f(x + \sqrt{2}t)dt \\ &= \begin{cases} 1 & \text{if } x \in [\epsilon, \frac{1}{2} - \epsilon) \\ -1 & \text{if } x \in [\frac{1}{2} + \epsilon, 1 - \epsilon). \end{cases} \end{aligned}$$

Let $B_j = [N_j, N_{j+1})$. Since the sequence $\{k\sqrt{2}\}$ is uniformly distributed in $[0, 1)$, we can choose the numbers N_j such that

$$\frac{\#\{k \in B_{2j}\colon \{k\sqrt{2}\} \in J_2\}}{N_{2j+1}} \geq |J_2| - \epsilon = \frac{1}{2} - \frac{2\sqrt{2}}{100}$$

and

$$\frac{\#\{k \in B_{2j-1}\colon \{k\sqrt{2}\} \in J_1\}}{N_{2j}} \geq |J_1| - \epsilon = \frac{1}{2} - \frac{2\sqrt{2}}{100}.$$

Then, for $x \in [0, 1/200]$,

$$\begin{aligned} P_{N_{2j+1}} f(x) &= \frac{1}{N_{2j+1}} \sum_{k<N_{2j}} I_{\epsilon_k} f(x + k\sqrt{2}) \\ &\quad + \frac{1}{N_{2j+1}} \sum_{\substack{k \in B_{2j} \\ \{k\sqrt{2}\} \in J_2}} I_{1/100} f(x + k\sqrt{2}) \\ &\quad + \frac{1}{N_{2j+1}} \sum_{\substack{k \in B_{2j} \\ \{k\sqrt{2}\} \notin J_2}} I_{1/\sqrt{2}} f(x + k\sqrt{2}) \\ &\leq \frac{1}{10} - \frac{\#\{k \in B_{2j}\colon \{k\sqrt{2}\} \in J_2\}}{N_{2j+1}} \\ &\leq \frac{1}{10} - \left(\frac{1}{2} - \frac{2\sqrt{2}}{100}\right) \leq -\frac{1}{4} \end{aligned}$$

and

$$\begin{aligned} P_{N_{2j}} f(x) &= \frac{1}{N_{2j}} \sum_{k<N_{2j-1}} I_{\epsilon_k} f(x + k\sqrt{2}) \\ &\quad + \frac{1}{N_{2j}} \sum_{\substack{k \in B_{2j-1} \\ \{k\sqrt{2}\} \in J_1}} I_{1/100} f(x + k\sqrt{2}) \\ &\quad + \frac{1}{N_{2j}} \sum_{\substack{k \in B_{2j-1} \\ \{k\sqrt{2}\} \notin J_1}} I_{1/\sqrt{2}} f(x + k\sqrt{2}) \\ &\geq -\frac{1}{10} + \frac{\#\{k \in B_{2j-1}\colon \{k\sqrt{2}\} \in J_1\}}{N_{2j}} \\ &\geq -\frac{1}{10} + \left(\frac{1}{2} - \frac{2\sqrt{2}}{100}\right) \geq \frac{1}{4}, \end{aligned}$$

showing that these averages diverge on a set of positive measure, even though $T_1 x = x + \sqrt{2}$ is an ergodic transformation.

4. Random sequences

In this section we require that the probability space (X, β, m) be a Lebesgue space in order to be able to use ergodic decompositions, and that $E = \{\epsilon_k\}$ be a stationary sequence of random variables defined on a probability space (Ω, C, P), taking values on $[0, 1]$.

Defining $\varphi_0 f(x) = f(x)$, the time averages now take the form

$$P_n f(x, w) = \frac{1}{n} \sum_{k=1}^{n} \varphi_{\epsilon_k(w)} f(T_k x)$$

where $x \in X$ and $w \in \Omega$.

The results of this section follow from the Return Times Theorem. Bourgain proved a first version of this theorem in [4], Bourgain, Furstenberg, Katznelson and Ornstein gave an alternate proof in [5], and Rudolph in [13] gave a proof through joinings and the formulation of the theorem that we present here.

Let S be a measure preserving transformation on a Lebesgue probability space (Ω, G, ν). The tuple (Ω, G, ν, S) is called a *dynamical system.* By abuse of notation, if (X, β, m) is a Lebesgue probability space and $\{T_t\}$ is a measurable flow of measure preserving transformations, then the tuple $(X, \beta, m, \{T_t\})$ is also called a *dynamical system.*

It will be clear from the context which definition of dynamical system we are using.

THEOREM 4.1. (RETURN TIMES THEOREM). *Let (Ω, G, ν, S) be a dynamical system, and let $g \in L^p(\nu)$. Then there is a subset $\Omega_g \subset \Omega$ of full measure so that for any other dynamical system (X, β, m, T) and $f \in L^q(m)$ $(1/p + 1/q = 1, 1 \leq p \leq \infty)$, if $w \in \Omega_g$,*

$$\lim_{k \to \infty} \frac{1}{n} \sum_{k=1}^{n} g(S^k w) f(T^k x)$$

converges for μ-a.e. $x \in X$.

Note 4.2. Given a sequence $\{\epsilon_k\}$ of stationary random variables defined on a probability space (Ω, C, P) and taking values in $[0, 1]$, they define a measure preserving one sided shift on the infinite product space $\Omega' = \times_{k=1}^{\infty}[0, 1]$ endowed with ν the measure induced by the distribution of the random variables. Let G be the completion of the Borel sets with respect to the Borel probability measure ν. Since Ω' is a complete separable metric space, a theorem of von Neumann [11] implies that (Ω', G, ν) is a Lebesgue space.

In what follows, we will not mention this infinite product construction but we are actually working in this space when we apply the Return Times Theorem.

LEMMA 4.3. *For any sequence $E = \{\epsilon_k\}$ of stationary random variables, there exists a set Ω_E of probability 1 such that for any dynamical system $(X, \beta, m, \{T_t\})$, if $w \in \Omega_E$, then $\lim_{n\to\infty} P_n f(x, w)$ exists for a.e. x and any $f \in L^\infty$.*

Proof. We will first show that there exists a set Ω_E such that for any dynamical system $(X, \beta, m, \{T_t\})$ where T_1 is ergodic, the averages $P_N f(x, w)$ converge for a.e. x.

Note that, by Proposition 2.4, for almost every x the map defined by F: $\epsilon \to \varphi_\epsilon(f)(x)$, $F(0) = f(x)$, is continuous and hence uniform continuous on $[0, 1]$.

Let $I_{n,i} = [i/n, (i+1)/n)$. Because the sequence of random variables $\{\epsilon_k\}_k$ is stationary and $f \in L^\infty$, it follows from the Return Times Theorem 4.1 that for each pair (i, n), $0 \le i < n$, there exists a set $\Omega_{i,n} \subset \Omega$ of probability 1 such that if $w \in \Omega_{i,n}$,

$$C_N^{n,i} f(x) = \frac{1}{N} \sum_{k=1}^{N} \chi_{I_{n,i}}(\epsilon_k(w))\varphi_{i/n} f(T_k x)$$

converges for a.e. $x \in X$ as $N \to \infty$, for any dynamical system (X, β, m, T_1). Then $\Omega_E = \bigcup_{n=1}^{\infty} \bigcup_{i=0}^{n-1} \Omega_{i,n}$ is a set of probability one.

Given $\delta > 0$ arbitrarily small, define

$$A_\delta^n = \left\{x\colon |\varphi_\epsilon f(x) - \varphi_{\epsilon'} f(x)| < \delta, \quad \text{if } |\epsilon - \epsilon'| \le \frac{1}{n}\right\}.$$

Because the uniform continuity of the map $\epsilon \to \varphi_\epsilon(f)(x)$, we can find an n_0 such that $m(A_\delta^n) > 1 - \delta$ for all $n \ge n_0$. Fixing $n \ge n_0$, if $f \in L^\infty$,

$$\begin{aligned}
&|P_N f(x, w) - P_M f(x, w)| \\
&\quad \le \left| P_N f(x, w) - \sum_{i=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} \chi_{I_{n,i}}(\epsilon_k(w))\varphi_{i/n} f(T_k x) \right| \\
&\quad\quad + \left| P_M f(x, w) - \sum_{i=0}^{n-1} \frac{1}{M} \sum_{k=1}^{M} \chi_{I_{n,i}}(\epsilon_k(w))\varphi_{i/n} f(T_k x) \right| \\
&\quad\quad + \sum_{i=0}^{n-1} |C_N^{n,i} f(x) - C_M^{n,i} f(x)| \\
&\quad = I_1 + I_2 + I_3.
\end{aligned}$$

To estimate the first two terms, observe that since T_1 is ergodic, for a.e. x, there exists $N_0 = N_0(x)$, such that for all $N \ge N_0$,

$$\frac{1}{N} \sum_{k=1}^{N} \chi_{(A_\delta^n)^c}(T_k x) < m((A_\delta^n)^c) + \delta.$$

Thus

$$\begin{aligned}
I_1 &= \sum_{i=0}^{n-1} \frac{1}{N} \sum_{k=1}^{N} \chi_{I_{n,i}}(\epsilon_k(w)) |\varphi_{\epsilon_k(w)} f(T_k x) - \varphi_{i/n} f(T_k x)| \\
&\leq \sum_{i=0}^{n-1} \frac{1}{N} \sum_{\substack{1 \leq k \leq N \\ T_k x \in A_\delta^n}} \chi_{I_{n,i}}(\epsilon_k(w)) \delta + \sum_{i=0}^{n-1} \frac{1}{N} \sum_{\substack{1 \leq k \leq N \\ T_k x \notin A_\delta^n}} \chi_{I_{n,i}}(\epsilon_k(w)) 2 \, \|f\|_\infty \\
&\leq \delta + 2\|f\|_\infty \, [m((A_\delta^n)^c) + \delta] \\
&\leq \delta \, [1 + 4\|f\|_\infty]
\end{aligned}$$

for all $N \geq N_0$.

The second term I_2 is handled in the same way, producing the same estimate provided that $M \geq N_0$.

To handle the the third term, recall that by the Return Times Theorem, for each $w \in \Omega_E$ and any pair (n, i), $0 \leq i < n$, $\{C_N^{n,i} f(x)\}_N$ is a Cauchy sequence, for a.e. x. Thus, for a.e. x, there exists $N_1 = N_1(x)$ such that if $N, M \geq N_1$, $|C_N^{n,i} f(x) - C_M^{n,i} f(x)| < \delta/n$, making $I_3 < \delta$.

Thus we have shown that for any $w \in \Omega_E$, given $\delta > 0$ for a.e. x there exist $\bar{N} = \max(N_0, N_1)$ such that for all $N, M \geq \bar{N}$, $|P_N f(x) - P_M f(x)| < \delta(3 + 8\|f\|_\infty)$.

Hence, for each $w \in \Omega_E$ and $f \in L^\infty$, the averages $P_N f(x, w)$ form a Cauchy sequence for a.e. $x \in X$. Therefore they converge.

The general case where T_1 may not be ergodic is readily obtained via the ergodic decomposition of the system (X, β, m, T_1).

If we have a flow for which T_1 is not ergodic, since (X, β, m) is a Lebesgue space, we can decompose the measure m into its ergodic components under T_1. Let (V, β', ν) be the factor of (X, β, m) defined by the sigma algebra of T_1–invariant sets of X. Then the collection of Borel measures $\{m_v\}_{v \in V}$ on the fibers has the property that (X, β, m_v, T_1) is an ergodic system for ν–a.e. $v \in V$, and

$$m(A) = \int_V m_v(A) d\nu \qquad \text{for all } A \in \beta.$$

(For a detailed explanation of decomposition over factor algebras see Rohlin [12], and for a simplified approach see Rudolph [14].)

Now, given $w \in \Omega_E$, let $C = \{x\colon \lim_{N \to \infty} P_N f(x, w) \text{ exists}\}$. Since (X, m_v, T_1) is an ergodic system for ν–a.e. $v \in V, m_v(C) = 1$ for ν–a.e. $v \in V$. Hence, $m(C) = 1$ also. $\square$

Applications of Hölder's inequality now yield convergence results in L^p.

Given any dynamical system $(X, \beta, m, \{T_t\})$, by Lemma 4.3 the set of functions $f \in L^p$ where $\lim_{n \to \infty} P_n f(x, w)$ exists for a.e. x, for all $w \in \Omega_E$, is dense in L^p. We need only to see that it is also closed.

Notice that in the above lemma, there was no integrability condition imposed on the stationary sequence $\{\epsilon_k\}$. For L^p convergence we will have to require some integrability condition on the reciprocal of ϵ_k, more precisely $1/\epsilon_k \in L^{q-1}$, with the convention

$$L^0 = \{\text{measurable functions on } \Omega\}.$$

THEOREM 4.4. *Let $E = \{\epsilon_k\}$ be a sequence of stationary random variables and p and q conjugate indexes, that is $1 \le p, q \le \infty$ and $1/p + 1/q = 1$. Assume that either φ has compact support and $\varphi \in L^q$, or $\phi \in L^1$. If $1/\epsilon_k \in L^{q-1}(\Omega)$, there is a set Ω_E of probability 1 such that if $w \in \Omega_E$, then for any dynamical system $(X, \beta, m, \{T_t\})$ and any $f \in L^p(X)$, $\lim_{n\to\infty} P_n f(x, w)$ exists for a.e. x.*

Proof. We will only prove the case $1 < p < \infty$. The case $p = 1$ is handled similarly. The case $p = \infty$ is nothing but Lemma 4.3.

Let Ω'_E be the set from Lemma 4.3, and

$$\Omega' = \left\{ w \in \Omega \colon \alpha(w) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} [1/\epsilon_k(w)]^{q-1} \text{ exists} \right\}.$$

Since $X_k = 1/\epsilon_k \in L^{q-1}(\Omega)$, Ω' is a set of probability 1. Define $\Omega_E = \Omega' \cap \Omega'_E$.

We will show that for any dynamical system, the subspace $H = \{f \in L^p\colon$ for all $w \in \Omega_E$, $\lim_{n\to\infty} P_n f(x, w)$ exists for a.e. $x\}$ is closed.

By the existence of the ergodic decomposition, we can assume that T_1 is ergodic. Let $f \in L^p$. Given $\delta > 0$, choose $g \in H$ such that $\|f - g\|_p < \delta$, and let $h = |f - g|$.

Case 1. φ has compact support and is in L^q.

Since the support of φ is compact, there exists L such that support of $\varphi \subset [-L, L]$. Then

$$\begin{aligned}
&|P_N(f-g)(x,w)| \\
&\quad \le \int_{|t|\le L} \left[\frac{1}{N}\sum_{k=1}^{N} |f-g|^p(T_{k+t}x)\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N} X_k^q(w)\varphi^q(t\,X_k(w))\right]^{1/q} dt \\
&\quad \le \left[\int_{|t|\le L} \frac{1}{N}\sum_{k=1}^{N} h^p(T_{k+t}x)\,dt\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N}\int_{|t|\le L} X_k^q(w)\varphi^q(t\,X_k(w))\,dt\right]^{1/q} \\
&\quad \le \left[\frac{1}{N}\sum_{k=1}^{N}\int_{|t|\le L} h^p(T_{k+t}x)\,dt\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N} X_k^{q-1}(w)\right]^{1/q} \|\varphi\|_q.
\end{aligned}$$

Then for each $w \in \Omega_E$,

$$\limsup_{N\to\infty} |P_N(f-g)(x,w)| \le (2L)^{1/p}\,\|h\|_p\,\alpha^{1/q}\,\|\varphi\|_q < (2L)^{1/p}\,\alpha^{1/q}\,\|\varphi\|_q\,\delta$$

for almost every x.

Case 2. $\phi(x) = \sup_{|y|\geq|x|} \varphi(y)$ is integrable.

$$
\begin{aligned}
&|P_N(f-g)(x,w)| \\
&\leq \int \left[\frac{1}{N}\sum_{k=1}^{N} |f-g|^p(T_{k+t}x)\,\varphi(tX_k)\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N} X_k^q(w)\varphi(t\,X_k(w))\right]^{1/q} dt \\
&\leq \left[\frac{1}{N}\sum_{k=1}^{N}\int h^p(T_{k+t}x)\,\phi(t)\,dt\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N}\int X_k^q(w)\phi(t\,X_k(w))\,dt\right]^{1/q} \\
&\leq \left[\frac{1}{N}\sum_{k=1}^{N}\int h^p(T_{k+t}x)\,\phi(t)\,dt\right]^{1/p} \left[\frac{1}{N}\sum_{k=1}^{N} X_k^{q-1}(w)\right]^{1/q} \|\phi\|_1^{1/q}
\end{aligned}
$$

Then, for each $w \in \Omega_E$,

$$\limsup_{N\to\infty} |P_N(f-g)(x)| \leq \|h\|_p\,\|\phi\|_1\,\alpha^{1/q} < \delta\,\|\phi\|_1\,\alpha^{1/q}$$

for almost every x.

From the estimates in either case it follows that if $w \in \Omega_E$ and $\delta > 0$,

$$\limsup_{N\to\infty} |P_N(f-g)(x,w)| \leq C(p,w)\,\delta \tag{2}$$

for almost every x. This is enough to show that the sequence $P_n(f)(x,w)$ is Cauchy. Indeed,

$$
\begin{aligned}
|P_n(f)(x,w) - P_m(f)(x,w)| \leq\ & |P_n(f-g)(x,w)| + |P_m(f-g)(x,w)| \\
& + |P_n(g)(x,w) - P_m(g)(x,w)|.
\end{aligned}
$$

Estimate 2 shows that for each $w \in \Omega_E$, there exist N_1 such that if $n, m \geq N_1$, the first two terms are $< C(p,w)\delta$ for almost every x. Since $g \in H$, if $w \in \Omega_E$, the sequence $P_n(g)(x,w)$ is Cauchy for a.e. x. Hence there is N_2 such that if $n, m \geq N_2$, the third term is smaller than δ. Hence, if $w \in \Omega_E$, for all $n, m \geq \max\{N_1, N_2\}$, $|P_n(f)(x,w) - P_m(f)(x,w)| < (1 + 2C(p,w))\delta$.

This proves that for every $w \in \Omega_E$, $P_n f(x,w)$ is a Cauchy sequence for a.e. x, and hence $f \in H$. □

References

[1] M. Akcoglu, A. Bellow, A. del Junco and R. Jones, *Divergence of averages obtained by sampling a flow*, Proc. Amer. Math. Soc. **118** (1993), 499–505.

[2] M. Akcoglu, A. del Junco and W. Lee, "A solution to a problem of Bellow" in *Almost everywhere convergence II*, A. Bellow and R. Jones ed., Academic Press, 1991, pp. 1–7.

[3] V. Bergelson, M. Boshernitzan and J. Bourgain, *Some results on non-linear recurrence*, J. Anal. Math. **62** (1994), 29–46.

[4] J. P. Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 5–41.
[5] J. P. Bourgain, H. Furstenberg, Y. Katznelson and D. S. Ornstein, *Return times of dynamical systems, Appendix*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 42–45.
[6] A. P. Calderón, *Ergodic theory and translation invariant operators*, Proc. Nat. Acad. Sci. **59** (1968), 349–353.
[7] K. L. Chung, *A course in probability theory*, 2nd ed., Academic Press, 1974.
[8] R. L. Jones and M. Wierdl, *Convergence and divergence of ergodic averages*, Ergodic Theory Dynamical Systems **14** (1994) no. 3, 515–535.
[9] ————, "Convergence of ergodic averages" in *Convergence in ergodic theory and probability*, Ohio State University Math. Research Inst. Pub. **5**, editors V. Bergelson, P. March and J. Rosenblatt, de Gruyter, New York, 1996, pp. 229–247.
[10] D. A. Lind, *Locally compact measure preserving flows*, Advances in Math. **15** (1975), 175–193.
[11] K. Petersen, *Ergodic theory*, Cambridge studies in advanced mathematics **2**, Cambridge University Press, 1983.
[12] V. A. Rohlin, *Selected topics from the metric theory of dynamical systems*, Amer. Math. Soc. Transl. **49** (1966), 171–240.
[13] D. Rudolph, *A joining's proof of Bourgain's return times theorem*, Ergodic Theory Dynamical Systems **14** (1994), 197–203.
[14] D. Rudolph, *Fundamentals of measurable dynamics. Ergodic theory on Lebesgue spaces*, Claredon Press, Oxford, 1990.
[15] C. Sadosky, *Interpolation of operators and singular integrals. An introduction to Harmonic Analysis*, Marcel Dekker, New York, 1979.
[16] A. Torchinsky, *Real variable methods in harmonic analysis*, Academic Press, 1986.

Department of Mathematics and Statistics, University at Albany, SUNY, Albany, NY 12222
reinhold@math.albany.edu