PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 87, Number 4, April 1983

NEW PROOFS FOR THE MAXIMAL ERGODIC THEOREM AND THE HARDY-LITTLEWOOD MAXIMAL THEOREM

ROGER L. JONES[1]

ABSTRACT. A new proof of the maximal ergodic theorem is presented. The same idea used in this proof is then used to show that the Hardy-Littlewood maximal function is weak type (1, 1).

1. In a probability space (X, Σ, m), let $T\colon X \to X$ be a measurable ($A \in \Sigma \Rightarrow T^{-1}A \in \Sigma$) measure preserving ($m(A) = m(T^{-1}A)$) ergodic ($TA = A \Rightarrow A = \varnothing$ or $A = X$) transformation.

THE MAXIMAL ERGODIC THEOREM. [1]. *Let $f \in L^1(X)$ and define $f^*(x) = \sup_n \sum_{k=0}^{n-1} f(T^k x)$. Then $\int_{\{f^* \geq 0\}} f(x)\, dm(x) \geq 0$.*

PROOF. By standard arguments it is enough to prove the result for $f \in L^\infty(X)$. Define

$$U_n = \left\{ x \mid \sum_{k=0}^{n-1} f(T^k x) \geq 0 \right\} \tag{1}$$

and let $E = \bigcup_{n=1}^{N} U_n$. We will show that $\int_E f(x)\, dm(x) \geq 0$. The general result follows by letting $N \to \infty$ in the definition of E.

Using the fact that T is measure preserving we have

$$\begin{aligned} \int_E f(x)\, dm(x) &= \int f(x)\chi_E(x)\, dm(x) \\ &= \frac{1}{L+1} \sum_{k=0}^{L} \int f(T^k x)\chi_E(T^k x)\, dm(x) \\ &= \int \frac{1}{L+1} \sum_{k=0}^{L} f(T^k x)\chi_E(T^k x)\, dm(x), \end{aligned} \tag{2}$$

where $L > N$ is arbitrary. (Think of L as very much larger than N.)

We will now show that for every x,

$$\frac{1}{L+1} \sum_{k=0}^{L} f(T^k x)\chi_E(T^k x) \geq \frac{-(N-1)\|f\|_\infty}{L+1}. \tag{3}$$

Received by the editors June 10, 1982.

1980 *Mathematics Subject Classification*. Primary 28D05, 47A35; Secondary 42A99.

Key words and phrases. Maximal ergodic theorem, maximal functions, Hardy-Littlewood maximal function.

[1]The research presented here was supported in part by a grant from the Faculty Research and Development Fund of the College of Liberal Arts and Sciences, DePaul University.

Since L is arbitrary, it follows that (2) is nonnegative, concluding the proof.

To obtain (3) we first establish the existence of a function $\tau(x) \leqslant N$ such that

$$\sum_{k=0}^{\tau(x)-1} f(T^k x)\chi_E(T^k x) \geqslant 0. \tag{4}$$

First note that

$$f(x)\chi_E(x) \geqslant f(x). \tag{5}$$

If $x \in E$ then (5) is obvious. If $x \notin E$ then $x \notin U_1$, which implies $f(x) < 0$, which implies (5) in this case also.

If $x \in E$ then $x \in U_n$ for some $n \leqslant N$. Using the definition of U_n and (5) we see that we can take $\tau(x) = n$. If $x \notin E$ then $\tau(x) = 1$ satisfies (4).

Now (3) follows. Sum the terms of

$$\sum_{k=0}^{L} f(T^k x)\chi_E(T^k x) \tag{6}$$

from the bottom index 0 to the index $\tau(x) - 1$. This is nonnegative by (4). Next sum the terms from the $\tau(x)$ to the index $\tau(T^{\tau(x)}x) - 1$. By (4) this sum is also nonnegative. Repeat this process until the end of the most recently considered partial sum has its top index greater than $L - N$.

The sum of the remaining (highest indexed) terms can be estimated by looking at the worst case. This sum has at most $N - 1$ terms, each at least $-\| f \|_\infty$, so that their sum, and hence the entire sum, exceeds $-(N - 1)\| f \|_\infty$. Dividing this estimate for (6) by $L + 1$ yields (3) and hence equation (1).

THE BIRKHOFF ERGODIC THEOREM. *For $f \in L^1(X)$ and a. e. x,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int f(x)\, dm(x). \tag{7}$$

PROOF. It is enough to prove the result for nonnegative $f \in L^\infty(X)$. The following use of the invariant sets U and V is due to Shields [2]. Define

$$U = \left\{ x \mid \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) > a \right\}$$

and

$$V = \left\{ x \mid \underline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) < b \right\}.$$

Assume $\int f(x)\, dm(x) = 1$, that $a > 1$ and $b < 1$. If (7) fails then there must exist (a, b) such that either $m(U) > 0$ or $m(V) > 0$. Since both U and V are invariant this means either $m(U) = 1$ or $m(V) = 1$.

If $m(U) = 1$ then define $g = f - a$. Let $U_n = \{x \mid \sum_{k=0}^{n-1} g(T^k x) \geqslant 0\}$. Then $U \subseteq \bigcup_{n=1}^{\infty} U_n$. Take N so large that $m(\bigcup_{n=1}^{N} U_n) > 1/a$, and let $E = \bigcup_{n=1}^{N} U_n$. By the maximal lemma, $\int_E g \geqslant 0$ which implies $\int_E f \geqslant am(E) > a(1/a) = 1 = \int f$, a contradiction.

If $m(V) = 1$ then the above argument can be repeated on the function $g = b - f$. At the end we arrive at $bm(E) \geqslant \int_E f$, where for any ε and δ we can have $m(E) > 1 - \varepsilon$ and $\int_E f > 1 - \delta$. Consequently we have $b(1 - \varepsilon) > 1 - \delta$ which is impossible since ε and δ can be as small as desired.

2. The methods used in §1 can be used to study the Hardy-Littlewood maximal function. We begin with a simple discrete version.

LEMMA. *Let $\{a_n\}_{n=1}^{N-1}$ be a sequence of nonnegative real numbers. Define $a_n^* = \sup_k (1/k)\sum_{j=0}^{k-1} a_{n+j}$, then*

$$\#\{n \mid a_n^* \geqslant \lambda\} \leqslant \frac{1}{\lambda}\sum_{n=1}^{N-1} a_n. \tag{8}$$

PROOF. Let $b_n = a_n - \lambda$, and let $E = \{a_n^* \geqslant \lambda\} = \{b_n^* \geqslant 0\}$. As before note that $b_n\chi_E(n) \geqslant b_n$. We now show that a function $\tau(n)$ exists as before. If $n \in E$ then there exists k such that $\sum_{j=0}^{k-1} b_{n+j} \geqslant 0$, which implies $\sum_{j=0}^{k-1} b_{n+j}\chi_E(n+j) \geqslant 0$, so we can set $\tau(n) = k$. If $n \notin E$ set $\tau(n) = 1$. The conclusion is that for every n

$$\sum_{k=0}^{\tau(n)-1} b_{n+k}\chi_E(n+k) \geqslant 0. \tag{9}$$

Let $n = 1$ and apply (9). Let $n = \tau(1) + 1$ and reapply (9). Continue until $N - 1$ is reached. Thus $\sum_{n=1}^{N-1} b_n\chi_E(n) \geqslant 0$, which can be rewritten as

$$\sum_{n=1}^{N-1} a_n\chi_E(n) \geqslant \lambda\#\{n \mid n \in E\}.$$

The lemma follows since $\{a_n\}$ is nonnegative.

THEOREM [3]. *For each $f \in L^1[0,1]$ define $f^*(x) = \sup_h (1/h)\int_x^{x+h} |f(t)|\,dt$, then $m\{f^* > \lambda\} \leqslant 1/\lambda \int |f(t)|\,dt$.*

PROOF. The idea is to reduce the problem to considering a finite number of points, and then apply the above lemma. (The theorem could be proved directly, just as in the proof of the lemma, if one is willing to use the axiom of choice.) Assume $f \geqslant 0$. Define $f_N(x,k) = N\int_{x+(k-1)/N}^{x+k/N} f(t)\,dt$ for $x \in (0, 1/N)$ and $0 < k < N$. Then

$$f_N^*(x,k) = \sup_n \frac{1}{n}\sum_{j=0}^{n-1} f_N(x, k+j) = \sup_n \frac{N}{n}\int_{x+(k-1)/N}^{x+(k-1)/N+n/N} f(t)\,dt.$$

Define $f_N^*(x) = \sup_n (1/(n/N))\int_x^{x+n/N} f(t)\,dt$. Then we have

$$\begin{aligned} m\{f_N^*(x) \geqslant \lambda\} &= \int_0^{1/N}\sum_{k=1}^{N-1} \chi(x)_{\{f_N^*(x+k-1/N)\geqslant\lambda\}}\,dx \\ &= \int_0^{1/N}\sum_{k=1}^{N-1} \chi(x)_{\{f_N^*(x,k)\geqslant\lambda\}}\,dx = \int_0^{1/N} \#\{k \mid f_n^*(x,k) \geqslant \lambda\}\,dx \\ &\leqslant \int_0^{1/N}\frac{1}{\lambda}\sum_{k=1}^{N-1} f_N(x,k)\,dx \leqslant \frac{1}{\lambda}\int_0^{1/N} N\int_x^{x+(N-1)/N} f(t)\,dt\,dx \\ &\leqslant \frac{1}{\lambda}\int_0^{1/N} N\int_0^1 f(t)\,dt\,dx \leqslant \frac{1}{\lambda}\int_0^1 f(t)\,dt. \end{aligned} \tag{10}$$

Since the estimate (10) is independent of N, the inequality holds with $\{f_N^* > \lambda\}$ replaced by $\{f^* > \lambda\}$, concluding the proof.

REMARK. The present paper was motivated by the work of Shields [2]. In his paper, Shields uses the fact that if $\{x \mid \overline{\lim}(1/n)\Sigma_{k=0}^{n-1} f(T_x^k) > \int f\}$ has positive measure then it has measure 1. As a consequence, most of the orbit of x is spent in "blocks" where some average of less than N terms becomes too large. He then must estimate the amount of time spent outside these "blocks". The replacement of $f(x)$ by $f(x)\chi_E(x)$ results in the entire orbit of x being spent in such blocks. Thus there is no need to estimate the time spent outside the blocks. In addition it is not necessary that E have large measure, which makes it possible to prove the maximal lemma as well as the limit theorem.

REFERENCES

1. P. R. Halmos, *Lectures on ergodic theory*, Japan Math. Soc., Kenkyusha Printing Co., Ltd., Tokyo, 1956.

2. P. C. Shields, *A simple, direct proof of Birkhoff's ergodic theorem*, preprint.

3. A. Zygmund, *Trigonometric series*, 2nd ed., Vol. 1, Cambridge Univ. Press, Cambridge, 1968, pp. 29–33.

DEPARTMENT OF MATHEMATICS, DEPAUL UNIVERSITY, 2323 N. SEMINARY, CHICAGO, ILLINOIS 60639