Z. Wahrscheinlichkeitstheorie 3, 75—88 (1964)

Maximal Inequalities as Necessary Conditions for Almost Everywhere Convergence*

By

D. L. Burkholder

Introduction

Consider the inequality

$$\mu(T^* f > \lambda) \leqq K \int_\Omega |f|^p \, d\mu / \lambda^p .$$

Here $(\Omega, \mathfrak{A}, \mu)$ is a σ-finite positive measure space, $1 \leqq p < \infty$, $f \in L_p(\Omega, \mathfrak{A}, \mu)$, $T^* f(\omega) = \sup_{1 \leqq n < \infty} |T_n f(\omega)|$ where each T_n is a bounded linear operator in L_p, $\lambda > 0$, and K is a real number not depending on f or λ.

Such an inequality, or one similar, is often the central part of proofs establishing the almost everywhere convergence of $\{T_n f\}$ for every f in L_p. The inequality implies, by letting $\lambda \to \infty$, that $T^* f < \infty$ almost everywhere for all f in L_p. Thus, by the Banach convergence theorem (see, for example, page 332 of [*6*]), the almost everywhere convergence of $\{T_n f\}$ for every f in L_p follows from the almost everywhere convergence of $\{T_n f\}$ for f in a dense subset of L_p. Convergence in a dense subset is often rather easy to establish. It is not uncommon, however, that proving the above sort of inequality, the maximal inequality, is genuinely difficult. Can one know in advance whether the maximal inequality approach to proving almost everywhere convergence is plausible? Conceivably, almost everywhere convergence could hold for the particular problem in which one is interested without a maximal inequality of the above form holding. When is such an inequality a necessary condition for almost everywhere convergence?

A. P. Calderón ([*12*], II, page 165) and E. M. Stein [*10*] have obtained important results in this direction for operators arising in Fourier analysis. For Stein, who generalizes Calderón's result considerably, Ω is the homogeneous space of a compact group and each T_n commutes with translations. However, many convergence problems in analysis, for example, those most often encountered in probability theory and ergodic theory, do not have this kind of setting.

In Section 1 of this paper we pose the necessity question somewhat more generally for arbitrary sequences of measurable functions. We consider a set $\mathscr{C}$ of such sequences and show that if each sequence in $\mathscr{C}$ converges almost everywhere (actually less is needed) and $\mathscr{C}$ satisfies one other condition, then $\mathscr{C}$ satisfies a maximal inequality (Theorems 1 and 2). These results apply to many problems in ergodic theory, probability theory, orthogonal series, and the like. Some of

* This research was supported by the National Science Foundation under grant G 21507.

these applications are discussed in the succeeding sections. The key requirement on $\mathscr{C}$ in Theorem 1 is that $\mathscr{C}$ be *stochastically convex*. This condition, defined in Section 1, is usually quite easy to check in the applications and implies that $\mathscr{C}$ is not too small.

Throughout the paper the same symbol is used for a measurable function and for the equivalence class (of measurable functions any two of which are equal almost everywhere) containing it. Expressions involving equality and inequality signs are sometimes to be interpreted as holding almost everywhere. Also, $\int f$ is occasionally used to denote the integral over Ω of f relative to the measure μ.

1. A basic question

Let $(\Omega, \mathfrak{A}, \mu)$ be a positive measure space with $\mu(\Omega) = 1$. Let $\mathscr{D}$ be the collection of all sequences $f = (f_1, f_2, \ldots)$ with each f_n an $\mathfrak{A}$-measurable function from Ω into the complex numbers. For $f = (f_1, f_2, \ldots)$ in $\mathscr{D}$ define f^* by $f^*(\omega) = \sup_{1 \leqq n < \infty} |f_n(\omega)|$, $\omega \in \Omega$. Let $\mathscr{C} \subset \mathscr{D}$ and $0 < p < \infty$.

Question. *What conditions on $\mathscr{C}$ assure the existence of a real number K satisfying*

$$\mu(f^* > \lambda) \leqq K/\lambda^p, \quad \lambda > 0, \quad f \in \mathscr{C}? \tag{1}$$

In particular, under what conditions on $\mathscr{C}$ does the almost everywhere convergence of each sequence in $\mathscr{C}$ imply the existence of such a K?

Note that the right hand side of (1) does not depend on f. This usually causes no difficulty in the applications and can often be accomplished by demanding that if $f = (f_1, f_2, \ldots) \in \mathscr{C}$, then $\int |f_1|^p \leqq 1$ or some other similar condition. Also, the condition $\mu(\Omega) = 1$ can often be dropped in the applications.

If f and g belong to $\mathscr{D}$ write $f \sim g$ if f and g have the same distribution, that is, if $\int \varphi(f) = \int \varphi(g)$ for all bounded Baire functions φ on the obvious product space. Clearly, if f and g belong to $\mathscr{D}$ and $f \sim g$, then $f^* \sim g^*$ where the notation is to denote again that f^* and g^* have the same distribution.

We shall say that $\mathscr{C}$ is *stochastically convex* if the following condition is satisfied: Each term of each sequence in $\mathscr{C}$ is nonnegative almost everywhere, and if $f_k = (f_{k1}, f_{k2}, \ldots) \in \mathscr{C}$, $k = 1, 2, \ldots$, then there are sequences $g_k = (g_{k1}, g_{k2}, \ldots) \in \mathscr{D}$, $k = 1, 2, \ldots$, such that

(i) the g_k's are (stochastically) independent,

(ii) $f_k \sim g_k$, $k = 1, 2, \ldots$,

(iii) if $\{a_k\}$ is a nonnegative number sequence with $\sum_{k=1}^{\infty} a_k = 1$, then there is an $h \in \mathscr{C}$ such that $h \sim \{\sum_{k=1}^{\infty} a_k g_{kn}\}$.

Note that (i) is equivalent to saying that the rows of the matrix (g_{kn}) are independent.

The *finiteness condition* is satisfied if $\mu(f^* < \infty) > 0$, $f \in \mathscr{C}$. Clearly, if each $f \in \mathscr{C}$ converges almost everywhere to a finite limit then the finiteness condition is satisfied.

Theorem 1. *Suppose that $\mathscr{C}$ is stochastically convex and satisfies the finiteness condition. Then there is a real number K such that*

$$\mu(f^* > \lambda) \leqq K/\lambda, \quad \lambda > 0, \quad f \in \mathscr{C}.$$

The convexity condition may be modified in various ways. One particular modification leads to

Theorem 2. *Let $0 < p < \infty$. Suppose that $\mathscr{C}$ satisfies the finiteness condition and condition C_p stated below. Then there is a real number K such that*

$$\mu(f^* > \lambda) \leqq K/\lambda^p, \quad \lambda > 0, \quad f \in \mathscr{C}.$$

Condition C_p: If $f_k = (f_{k1}, f_{k2}, \ldots) \in \mathscr{C}$, $k = 1, 2, \ldots$, then there are sequences $g_k = (g_{k1}, g_{k2}, \ldots) \in \mathscr{D}$, $k = 1, 2, \ldots$, such that

(i) *the g_k's are independent,*

(ii) $f_k^* \sim g_k^*$, $\quad k = 1, 2, \ldots,$

(iii) *if $\{a_k\}$ is a real number sequence satisfying $\sum_{k=1}^{\infty} |a_k|^p = 1$, then the series $\sum_{k=1}^{\infty} a_k g_{kn}$ converges almost everywhere, $n = 1, 2, \ldots$, and there is an $h \in \mathscr{C}$ and a $\theta > 0$ such that $h \sim \{\theta \sum_{k=1}^{\infty} a_k g_{kn}\}$.*

Proof of Theorem 1. We have to show that the function M defined by

$$M(\lambda) = \sup_{f \in \mathscr{C}} \lambda \mu(f^* > \lambda)$$

is bounded on $(0, \infty)$. Suppose this is not true. Then $\limsup_{\lambda \to \infty} M(\lambda) = \infty$ and it is easy to see that this implies the existence of positive number sequences $\{a_k\}$ and $\{\lambda_k\}$ satisfying $\lim_{k \to \infty} \lambda_k = \infty$, $\sum_{k=1}^{\infty} a_k = 1$, and $\sum_{k=1}^{\infty} (a_k/\lambda_k)\, M(\lambda_k/a_k) = \infty$. Let $f_k \in \mathscr{C}$ satisfy

$$\mu(f_k^* > \lambda_k/a_k) > (a_k/\lambda_k)\, M(\lambda_k/a_k) - 1/2^k,$$

$k = 1, 2, \ldots$, and let $g_1, g_2, \ldots$ be as in the stochastic convexity condition. Then $\mu(g_k^* > \lambda_k/a_k) = \mu(f_k^* > \lambda_k/a_k)$ and therefore $\sum_{k=1}^{\infty} \mu(g_k^* > \lambda_k/a_k) = \infty$. Since the sets involved are independent, we have by Borel that for almost all ω, $g_k^*(\omega) > \lambda_k/a_k$ for infinitely many positive integers k, which implies that almost everywhere

$$\limsup_{k \to \infty} a_k g_k^* = \infty. \tag{2}$$

Let $h \in \mathscr{C}$, $h \sim \{\sum_{k=1}^{\infty} a_k g_{kn}\}$. Then by nonnegativity, we have that almost everywhere

$$\sup_n \sum_{k=1}^{\infty} a_k g_{kn} \geqq \sup_n \sup_k a_k g_{kn} = \sup_k a_k g_k^* = \infty ,$$

which implies that $h^* = \infty$ almost everywhere, a contradiction of the finiteness condition. Thus, M is bounded and the theorem is proved.

In the proof of Theorem 2 we shall use a typical result about Rademacher functions. For real t, let

$$r_1(t) = 1 \quad \text{if} \quad 0 \leqq t \leqq 1/2, = -1 \quad \text{if} \quad 1/2 < t < 1,$$

$$r_1(t+1) = r_1(t), \quad \text{and} \quad r_{n+1}(t) = r_1(2^n t), \quad n = 1, 2, \ldots .$$

Lemma 1. *Suppose that for all t in a set of positive Lebesgue measure the series $\sum_{k=1}^{\infty} a_{nk} r_k(t)$ converges, $n = 1, 2, \ldots$, and $\sup_{1 \leqq n < \infty} \left| \sum_{k=1}^{\infty} a_{nk} r_k(t) \right| < \infty$. Then*

$$\limsup_{k \to \infty} \sup_n |a_{nk}| < \infty .$$

This is an immediate consequence of a fact about Rademacher series mentioned, for example, in STEIN ([*10*], Lemma 2).

Proof of Theorem 2. The proof is similar to the proof of Theorem 1 up to and including the establishment of (2). The definition of M is changed to

$$M(\lambda) = \sup_{f \in \mathscr{C}} \lambda^p \mu(f^* > \lambda) ,$$

the unboundedness of M implies the existence of positive number sequences $\{a_k\}$ and $\{\lambda_k\}$ satisfying $\lim_{k \to \infty} \lambda_k = \infty$, $\sum_{k=1}^{\infty} a_k^p = 1$, $\sum_{k=1}^{\infty} (a_k/\lambda_k)^p M(\lambda_k/a_k) = \infty$, and much as before we have that almost everywhere

$$\limsup_{k \to \infty} a_k g_k^* = \infty .$$

By condition C_p and the finiteness condition we have that for each t, the series $\sum_{k=1}^{\infty} r_k(t) a_k g_{kn}$ converges almost everywhere, $n = 1, 2, \ldots$, and

$$\sup_n \left| \sum_{k=1}^{\infty} r_k(t) a_k g_{kn} \right| < \infty$$

on a set of positive measure. By Fubini's theorem, for all ω in a set of positive μ measure, the conditions of Lemma 1 are satisfied by $a_{nk} = a_k g_{kn}(\omega)$, hence for such ω,

$$\limsup_{k \to \infty} a_k g_k^*(\omega) = \limsup_{k \to \infty} \sup_n |a_k g_{kn}(\omega)| < \infty ,$$

giving a contradiction. This completes the proof.

The following is helpful in some of the applications. Given $\mathscr{C} \subset \mathscr{D}$, define $\overline{\mathscr{C}}$ as follows: The sequence $\bar{f} = (\bar{f}_1, \bar{f}_2, \ldots) \in \overline{\mathscr{C}}$ if and only if there is an $f = (f_1, f_2, \ldots) \in \mathscr{C}$ such that

$$\bar{f}_n = (f_1 + \cdots + f_n)/n, \quad n = 1, 2, \ldots.$$

Theorem 3. *If $\mathscr{C}$ is stochastically convex, then $\overline{\mathscr{C}}$ is stochastically convex. If $\mathscr{C}$ satisfies condition C_p', then $\overline{\mathscr{C}}$ satisfies condition C_p'.*

Here C_p' is the condition C_p modified by replacing (ii) of that condition by (ii) of the stochastic convexity condition.

The proof of Theorem 3 is clear.

2. Stochastic convexity — examples

Throughout this section let $(\Omega, \mathfrak{A}, \mu)$ be Lebesgue measure on the Borel sets of the unit interval. Additional examples of stochastic convexity appear in later sections.

Strictly stationary processes. Let $\mathscr{C}$ be the set of all strictly stationary sequences $f = (f_1, f_2, \ldots)$ with $0 \leqq f_1$ and $\int f_1 \leqq 1$. Recall that f is strictly stationary if $f \in \mathscr{D}$ and $(f_1, \ldots, f_n) \sim (f_{m+1}, \ldots, f_{m+n})$ for all m, n. Then $\mathscr{C}$ hence also $\overline{\mathscr{C}}$ is stochastically convex. To see this let I be the identity function on Ω and let $s_1, s_2, \ldots$ be independent functions on Ω satisfying $s_k \sim I, k = 1, 2, \ldots$. The nonnegativity condition is satisfied. Let $f_k = (f_{k1}, f_{k2}, \ldots) \in \mathscr{C}$, $k = 1, 2, \ldots$. Define sequences g_k by $g_{kn} = f_{kn}(s_k)$. Clearly, (i) and (ii) of the stochastic convexity condition are satisfied. To see that (iii) is also satisfied let $\{a_k\}$ be a nonnegative number sequence with $\sum_{k=1}^{\infty} a_k = 1$. Then, for each n, $\int \sum_{k=1}^{\infty} a_k g_{kn} = \sum_{k=1}^{\infty} a_k \int g_{kn} \leqq 1$. Thus, there is a sequence $h = (h_1, h_2, \ldots) \in \mathscr{D}$ such that $0 \leqq h_1$, $\int h_1 \leqq 1$, and $h_n = \sum_{k=1}^{\infty} a_k g_{kn}$ almost everywhere, $n = 1, 2, \ldots$. The strict stationarity of h follows from the fact that $(G_1, G_2, \ldots, G_n) \sim (G_{m+1}, \ldots, G_{m+n})$ for all m, n where $G_n = (g_{1n}, g_{2n}, \ldots)$. Thus $h \in \mathscr{C}$ implying that $\mathscr{C}$ is stochastically convex. The stochastic convexity of $\overline{\mathscr{C}}$ follows by Theorem 3.

Martingales. Here let $\mathscr{C}$ be the set of all $f = (f_1, f_2, \ldots)$ in $\mathscr{D}$ satisfying $0 \leqq f_n$, $\int f_n \leqq 1$, and $E(f_{n+1} | f_1, \ldots, f_n) \geqq f_n$, $n = 1, 2, \ldots$. Each f in $\mathscr{C}$ is a submartingale (= semimartingale [4]). It is easy to see that $\mathscr{C}$ is stochastically convex. One may proceed exactly as in the above example until h is obtained. The only nontrivial step is to verify that h is a submartingale. Using the fact that $f_k \in \mathscr{C}$, $f_k \sim g_k$ implies that g_k is a submartingale, we have that

$$\begin{aligned} &E(h_{n+1} | (g_{k1}, \ldots, g_{kn}), \quad k = 1, 2, \ldots) \\ &= \sum_{k=1}^{\infty} a_k E(g_{k,n+1} | g_{k1}, \ldots, g_{kn}) \\ &\geqq \sum_{k=1}^{\infty} a_k g_{kn} = h_n, \end{aligned}$$

and the desired result follows by operating on both ends of this inequality with the conditional expectation $E(\cdot | h_1, \ldots, h_n)$.

Similar results can be obtained for martingales, supermartingales, backward martingales, and the like.

Alternating processes. Here let $f = (f_0, f_1, \ldots)$ belong to $\mathscr{C}$ if and only if $0 \leqq f_0$, $\int f_0 \leqq 1$, and there are two conditional expectations T_1 and T_2 such that $T_1 f_{2n} = f_{2n+1}$ and $T_2 f_{2n+1} = f_{2n+2}$, $n = 0, 1, \ldots$. It follows that $f \in \mathscr{C}$ if and only if $0 \leqq f_0$, $\int f_0 \leqq 1$, and

$$(3) \qquad \begin{aligned} E(f_{2n} \mid f_1, f_3, \ldots) &= f_{2n+1} \\ E(f_{2n+1} \mid f_2, f_4, \ldots) &= f_{2n+2}, \end{aligned}$$

$n = 0, 1, \ldots$. To show that $\mathscr{C}$ is stochastically convex one may proceed again as in the first example to obtain a sequence h which must be shown to be in $\mathscr{C}$. Since $f_k \in \mathscr{C}$, $f_k \sim g_k$ implies that g_k is also an alternating process (satisfies (3)), we have that

$$\begin{aligned} &E(h_{2n} \mid (g_{k1}, g_{k3}, \ldots), \quad k = 1, 2, \ldots) \\ &= \sum_{k=1}^{\infty} a_k E(g_{k,2n} \mid g_{k1}, g_{k3}, \ldots) \\ &= \sum_{k=1}^{\infty} a_k g_{k,\, 2n+1} = h_{2n+1}, \end{aligned}$$

and the first part of (3) follows by operating on both ends of this relation with the conditional expectation $E(\cdot \mid h_1, h_3, \ldots)$. The proof that h satisfies the second part of (3) is similar.

Remark. By the ergodic theorem for strictly stationary processes, each sequence in $\overline{\mathscr{C}}$ of the first example converges almost everywhere. The same is true for $\mathscr{C}$ of the second example by DOOB's convergence theorem for submartingales [*4*]. Also, maximal inequalities are known to hold for both cases. However, the third example is different in this regard. At present, neither the almost everywhere convergence question nor the maximal inequality question is settled. However, since stochastic convexity is satisfied here, the two questions are equivalent by Theorem 1 and the Banach convergence theorem. In the next section we shall investigate this problem further.

3. Applications to a problem in operator ergodic theory

Let $(\Omega, \mathfrak{A}, \mu)$ be a positive measure space and T a linear positive definite self-adjoint operator in $L_1(\Omega, \mathfrak{A}, \mu)$ with norm $\|T\|_1 \leqq 1$. Positive definiteness is to mean here that the inner product (Tf, f) is nonnegative for all f in $L_1 \cap L_\infty$; self-adjointness, that $(Tf, g) = (f, Tg)$ for all f and g in $L_1 \cap L_\infty$. Our assumptions imply that T has a bounded linear extension in L_p and that this extension, also denoted by T, does not increase the L_p norm of any function in L_p, $1 < p < \infty$. It is known that if $1 < p < \infty$ and $f \in L_p$, then $\{T^n f\}$ converges almost everywhere. The $p = 2$ case of this proposition is due to CHOW and the present author ([*3*], Theorem 2 and the remark following Lemma 2); the general case is due to STEIN [*11*]. For a different proof of the major part of the general result, see ROTA [*8*]. If $f \in L_1$ does $\{T^n f\}$ necessarily converge almost everywhere? At the present time, this question still seems to be open. A related almost everywhere convergence result due to Rota [8] does not extend to L_1 [2]. Some information about the L_1 case is contained in the following theorem. This information may help to lead to an answer to the above question.

In the following let $T^*f = \sup_{0 \leq n < \infty} |T^n f|$.

Theorem 4. *The following statements are equivalent:*

(i) *If* $(\Omega, \mathfrak{A}, \mu)$ *is a positive measure space and* T *is a linear positive definite self-adjoint operator in* $L_1(\Omega, \mathfrak{A}, \mu)$ *with norm* $\|T\|_1 \leq 1$, *then* $\{T^n f\}$ *converges almost everywhere for each* $f \in L_1$.

(ii) *If* $(\Omega, \mathfrak{A}, \mu)$ *is Lebesgue measure on the Borel sets of the unit interval,* T_1 *and* T_2 *conditional expectations, and* $T = T_1 T_2 T_1$, *then* $\{T^n f\}$ *converges almost everywhere for each* $f \in L_1(\Omega, \mathfrak{A}, \mu)$.

(iii) *There is a real number* K *such that if* r *is a positive integer,* $\Omega = \{1, 2, \ldots, r\}$, μ *is uniform probability over* Ω, *and* $Tf(j) = \sum_{k=1}^{r} p_{jk} f(k)$, *where* (p_{jk}) *is a symmetric stochastic matrix of order* r, *then*

$$\mu(T^*f > \lambda) \leq K \int_\Omega |f| \, d\mu/\lambda, \quad \lambda > 0, \quad f \in L_1.$$

(iv) *There is a real number* K *such that if* $(\Omega, \mathfrak{A}, \mu)$ *is a positive measure space and* T *is a linear self-adjoint operator in* $L_1(\Omega, \mathfrak{A}, \mu)$ *with norm* $\|T\|_1 \leq 1$, *then*

$$\mu(T^*f > \lambda) \leq K \int_\Omega |f| \, d\mu/\lambda, \quad \lambda > 0, \quad f \in L_1.$$

Proof. (i) $\Rightarrow$ (ii): The operator $T = T_1 T_2 T_1$ described in (ii) satisfies the condition on T in (i).

(ii) $\Rightarrow$ (iii): Suppose (ii) holds. Then the set $\mathscr{C}$ described in the third example of Section 2 satisfies the finiteness condition since if $f = (f_0, f_1, \ldots) \in \mathscr{C}$ then there are conditional expectations T_1 and T_2 such that $\{f_{2n+1}\} = \{(T_1 T_2 T_1)^n f_1\}$ and $\{f_{2n+2}\} = \{(T_2 T_1 T_2)^n f_2\}$. Therefore, since $\mathscr{C}$ satisfies the stochastic convexity condition, we have by Theorem 1 that $\mathscr{C}$ satisfies a maximal inequality of the type described in Theorem 1. This implies the existence of a real number K such that if T_1 and T_2 are conditional expectations, then

$$\mu\left(\sup_{0 \leq n < \infty} |(T_1 T_2 T_1)^n f| > \lambda\right) \leq K \int_\Omega |f| \, d\mu/\lambda, \quad \lambda > 0, \quad f \in L_1. \tag{4}$$

Now let T and (p_{jk}) be as in (iii). Partition the unit interval into disjoint connected sets B_{jk} such that $\mu(B_{jk}) = p_{jk}/r$, $j = 1, \ldots, r$; $k = 1, \ldots, r$. Let $A_j = \bigcup_{k=1}^{r} B_{jk}$, $B_k = \bigcup_{j=1}^{r} B_{jk}$. Then $\mu(A_j) = \mu(B_k) = 1/r$. Let $\mathfrak{A}_1$ be the smallest σ-field containing $\{A_1, \ldots, A_r\}$ and let $\mathfrak{A}_2$ be the smallest σ-field containing $\{B_1, \ldots, B_r\}$. Let $T_1 = E(\cdot \mid \mathfrak{A}_1)$ and $T_2 = E(\cdot \mid \mathfrak{A}_2)$. If f is a function on $\{1, \ldots, r\}$ let $g(\omega) = f(j)$ if $\omega \in A_j$. Then straightforward calculation shows that $T^2 f(j) = T_1 T_2 T_1 g(\omega)$ if $\omega \in A_j$, which implies more generally that $T^{2n} f(j) = (T_1 T_2 T_1)^n g(\omega)$ if $\omega \in A_j$. Note that $\{T^{2n} f\} \sim \{(T_1 T_2 T_1)^n g\}$. Clearly, (iii) follows from (4) using the fact that $T^*f \leq \sup_n |T^{2n} f| + \sup_n |T^{2n}(Tf)|$.

(iii) $\Rightarrow$ (iv): This is proved by standard approximation arguments. Let λ be a positive number and N a positive integer. Let T and f be as in (iv). It suffices to bound $\mu\ (\sup_{0 \leqq n \leqq N} |T^n f| > \lambda)$ appropriately.

Let $\Omega_k = \{\omega |\ \sup_{0 \leqq n \leqq N} |T^n f(\omega)| > 1/k\}$, U_k be the operation of multiplication by the characteristic function of Ω_k, $T_k = U_k T U_k$, and $f_k = U_k f$, $k = 1, 2, \ldots$. Then $\mu(\Omega_k) < \infty$, and the operator T_k, essentially acting in a finite measure space, satisfies the assumptions of (iv). Since, for $0 \leqq n \leqq N$, $U_k T^n f \rightarrow T^n f$ in L_1 norm as $k \rightarrow \infty$, we have by an induction argument that $T_k^n f_k \rightarrow T^n f$ in L_1 norm as $k \rightarrow \infty$, $0 \leqq n \leqq N$. This implies that

$$\mu\ (\sup_{0 \leqq n \leqq N} |T^n f| > \lambda) \leqq \sup_{1 \leqq k < \infty} \mu\ (\sup_{0 \leqq n \leqq N} |T_k^n f_k| > \lambda)\,. \tag{5}$$

On the other hand, $\int |f_k| \leqq \int |f|$ for all k. We need to consider our problem, therefore, only for finite measure spaces, and hence only for those satisfying $\mu(\Omega) = 1$.

Assume that $\mu(\Omega) = 1$ but let T, f, λ, N be as before. Let $\{B_1, B_2, \ldots\}$ be a countable collection of sets in $\mathfrak{A}$ such that $T^n f$ is measurable with respect to the smallest σ-field containing $\{B_1, B_2, \ldots\}$, $n = 0, 1, \ldots, N$. Let $\mathfrak{B}_k$ be the smallest σ-field containing $\{B_1, \ldots, B_k\}$, U_k the conditional expectation $E(\cdot \mid \mathfrak{B}_k)$, $T_k = U_k T U_k$, and $f_k = U_k f$, $k = 1, 2, \ldots$. The operator T_k satisfies the assumptions of (iv) and again we have (5) for this case. This time we conclude that our problem need be considered only for $\mathfrak{A}$ finite, hence for Ω finite. By appropriately splitting up the measure placed on each point of Ω, it is easy to see, again by an approximation argument, that equal measure may be placed on each point of Ω.

Therefore, assume that $(\Omega, \mathfrak{A}, \mu)$ is uniform probability over $\{1, 2, \ldots, r\}$ where r is a positive integer. Let T, f, λ, N be as before. Then $T^n f(j) = \sum_{k=1}^{r} a_{jk}^{(n)} f(k)$, $j = 1, \ldots, r$; $n = 1, 2, \ldots$, where $(a_{jk}^{(n)})$ is the n-th power of a matrix (a_{jk}) equal to its conjugate transpose and such that $\sum_{j=1}^{r} |a_{jk}| \leqq 1$, $k = 1, \ldots, r$. Let $p_{jk} = |a_{jk}|, j \neq k$, and define p_{jj} by $\sum_{k=1}^{r} p_{jk} = 1$. Then (p_{jk}) is a symmetric stochastic matrix satisfying

$$|T^n f(j)| \leqq \sum_{k=1}^{r} |a_{jk}^{(n)}|\,|f(k)| \leqq \sum_{k=1}^{r} p_{jk}^{(n)} |f(k)|\,,$$

which implies, assuming that (iii) holds, that

$$\mu(T^* f > \lambda) \leqq K \int_{\Omega} |f|\, d\mu / \lambda$$

where K is the constant of (iii). This completes the proof that (iii) $\Rightarrow$ (iv). (Note that a value of K working for (iii) works for (iv) and conversely.)

(iv) $\Rightarrow$ (i): Let T be as in (i). If $f \in L_1$, then the set where $T^* f > 0$ is the union of a countable number of sets of finite measure. Accordingly, one may suppose that $(\Omega, \mathfrak{A}, \mu)$ is σ-finite. We know that $\{T^n f\}$ converges almost everywhere for all f in a dense subset of L_1. If (iv) holds we have that if $f \in L_1$ then $T^* f < \infty$ almost everywhere by taking $\lambda \to \infty$ in the maximal inequality. Therefore, by the Banach convergence theorem, (iv) $\Rightarrow$ (i).

4. On the Hopf ergodic theorem and the Dunford-Schwartz ergodic theorem

A linear operator T in L_1 of a probability space $(\Omega, \mathfrak{A}, \mu)$ is *doubly stochastic* (confer [8]) if, for f in L_1, $Tf \geqq 0$ if $f \geqq 0$, $\int Tf = \int f$, and $T1 = 1$.

Let $(\Omega, \mathfrak{A}, \mu)$ be Lebesgue measure on the Borel sets of the unit interval. Let $\mathscr{C}$ be the set of all sequences $(f, Tf, T^2f, \ldots)$ where $0 \leq f$, $\int f \leq 1$, and T is doubly stochastic. (Note that T may change from sequence to sequence.)

Theorem 5. *The set $\mathscr{C}$ hence also $\overline{\mathscr{C}}$ is stochastically convex.*

The Hopf ergodic theorem [7] implies that each sequence in $\overline{\mathscr{C}}$ converges almost everywhere. Thus, by Theorems 1 and 5, a maximal inequality holds for $\overline{\mathscr{C}}$. By using arguments almost exactly the same as those used to prove the (iii) $\Rightarrow$ (iv) part of Theorem 4, one can show that this implies the following (confer [9]): There is a real number K such that if $(\Omega, \mathfrak{A}, \mu)$ is a positive measure space and T is a linear operator in $L_1(\Omega, \mathfrak{A}, \mu)$ such that $||T||_1 \leqq 1$ and $||T||_\infty \leqq 1$, then

$$\mu\,(\sup_{1 \leqq n < \infty} |\sum_{k=0}^{n-1} T^k f|/n > \lambda) \leqq K \int_\Omega |f|\, d\mu/\lambda\,,$$

$\lambda > 0$, $f \in L_1$, from which follows easily the almost everywhere convergence of each such sequence $\{\sum_{k=0}^{n-1} T^k f/n\}$, which is the content of the discrete one-parameter Dunford-Schwartz ergodic theorem [5]. Thus, the Dunford-Schwartz theorem follows from the more special Hopf theorem.

Again a maximal inequality turns out to be a necessary as well as sufficient condition for almost everywhere convergence. This would have been of particular interest during the period of eighteen years or more in which attempts at proving the almost everywhere convergence were being made.

Proof of Theorem 5. Suppose that $(f_k, T_k f_k, T_k^2 f_k, \ldots) \in \mathscr{C}$, $k = 1, 2, \ldots$. Let I be the identity function on Ω and let $s_1, s_2, \ldots$ be independent functions on Ω satisfying $s_k \sim I$, $k = 1, 2, \ldots$. Also, suppose that $\mathfrak{A}$ is the smallest σ-field with respect to which every s_k is measurable. Consider the set of all functions of the form

$$\sum_{j_1} \cdots \sum_{j_n} a(j_1, \ldots, j_n)\, \varphi_{1j_1}(s_1) \cdots \varphi_{nj_n}(s_n) \tag{6}$$

where n is a positive integer, the sum contains only a finite number of terms, the a's are scalars, the φ's are characteristic functions of sets in $\mathfrak{A}$, and $\sum_{j_i} \varphi_{ij_i} = 1$, $i = 1, \ldots, n$. This set, a linear manifold, is dense in $L_1(\Omega, \mathfrak{A}, \mu)$. At a function (6) in this set, define the value of an operator T by

$$\sum_{j_1} \cdots \sum_{j_n} a(j_1, \ldots, j_n)\, (T_1 \varphi_{1j_1})(s_1) \cdots (T_n \varphi_{nj_n})(s_n)\,.$$

Straightforward calculations show that this uniquely defines T and that T satisfies the conditions of double stochasticity on the set in question. Hence T can be extended to a doubly stochastic operator in L_1. Denoting the extension also by T, we see that $T^n[f_k(s_k)] = (T_k^n f_k)(s_k)$ for all k, n. Therefore

$$(f_k, T_k f_k, \ldots) \sim (f_k(s_k), \quad T[f_k(s_k)], \ldots), \quad k = 1, 2, \ldots,$$

the sequences on the right are independent, and if $\{a_k\}$ is a nonnegatve number sequence with $\sum_{k=1}^{\infty} a_k = 1$, then $\{T^n \sum_{k=1}^{\infty} a_k f_k(s_k)\}$ belongs to $\mathscr{C}$ and has the same distribution as $\{\sum_{k=1}^{\infty} a_k T^n[f_k(s_k)]\}$. Thus, $\mathscr{C}$ is stochastically convex. The stochastic convexity of $\overline{\mathscr{C}}$ follows by Theorem 3.

Remark. If in the definition of $\mathscr{C}$ it is also required that T be self-adjoint then again $\mathscr{C}$ is stochastically convex. This provides an alternative approach to some of the results of Section 3.

5. Orthogonal series of the Walsh-Paley type

Let $(\Omega, \mathfrak{A}, \mu)$ be a positive measure space with $\mu(\Omega) = 1$. Let $\varphi_0, \varphi_1, \ldots$ be independent real functions such that $\int \varphi_k = 0$ and $\int \varphi_k^2 = 1$, $k = 0, 1, \ldots$. Let $\psi_0 = 1$ and, if $n > 0$, let $\psi_n = \varphi_{n_1}\varphi_{n_2} \cdots \varphi_{n_k}$ where $n = 2^{n_1} + 2^{n_2} \cdots + 2^{n_k}$ and $n_1 > n_2 > \cdots > n_k \geqq 0$. Clearly, $\{\psi_n\}$ is an orthonormal sequence. If $\{\varphi_n\}$ is the sequence of Rademacher functions, then $\{\psi_n\}$ is the sequence of Walsh-Paley functions. However, we do not limit ourselves to this special case.

Theorem 6. *Let* $S_n = \sum_{k=0}^{n-1} a_k \psi_k$, $n = 1, 2, \ldots$, *where* $\{a_k\}$ *is a complex number sequence. Then* $\{S_{2^n}\}$ *is a martingale.*

Proof. We have that

$$S_{2^{n+1}} = S_{2^n} + \sum_{k=2^n}^{2^{n+1}-1} a_k \psi_k = S_{2^n} + \varphi_n \sum_{k=0}^{2^n-1} a_{2^n+k} \psi_k .$$

Operating on both sides with the conditional expectation $E(\cdot \mid \varphi_0, \ldots, \varphi_{n-1})$ gives $E(S_{2^{n+1}} \mid \varphi_0, \ldots, \varphi_{n-1}) = S_{2^n}$ which implies the desired result. We have used the fact that S_{2^n} and the sum multiplying φ_n are functions of $\varphi_0, \ldots, \varphi_{n-1}$, and that

$$E(\varphi_n \mid \varphi_0, \ldots, \varphi_{n-1}) = 0 . \tag{7}$$

Note that Theorem 6 remains true if the only assumption made about $\{\varphi_n\}$ is (7).

From now on, if f is a measurable function such that the inner products (f, ψ_k) exist, let $S_n f = \sum_{k=0}^{n-1} (f, \psi_k)\psi_k$ denote the n-th partial sum of the orthogonal expansion of f in terms of $\psi_0, \psi_1, \ldots$.

Corollary 1. *If* $f \in L_2(\Omega, \mathfrak{A}, \mu)$ *then* $\{S_{2^n} f\}$ *converges almost everywhere.*

Proof. We have that

$$\int |S_{2^n} f|^2 = \sum_{k=0}^{2^n-1} |(f, \psi_k)|^2 \leqq \int |f|^2 < \infty .$$

By Theorem 6 and the martingale convergence theorem [*4*], the desired result follows.

If A_k is a set containing exactly two numbers and $\mu(\varphi_k \in A_k) = 1$, we shall say that φ_k has a *two-point distribution*. If φ_k has a two-point distribution, $k = 0, 1, 2, \ldots$, then we shall say that $\{\varphi_k\}$ satisfies the *two-point condition*.

Theorem 7. *Suppose that $\{\varphi_k\}$ satisfies the two-point condition. If $1 \leqq p < \infty$ and $f \in L_p(\Omega, \mathfrak{A}, \mu)$, then*

$$S_{2^n} f = E(f \mid \varphi_0, \ldots, \varphi_{n-1}), \quad n = 1, 2, \ldots, \tag{8}$$

and $\{S_{2^n} f\}$ converges almost everywhere and in L_p norm. Suppose that, in addition, $\mathfrak{A}$ is the smallest σ-field with respect to which every φ_k is measurable. Then the above sequence converges to f and $\{\psi_n\}$ is a complete orthonormal system.

Proof. It follows from the assumptions on $\{\varphi_k\}$ that there exist numbers a_{jk}, b_{jk} such that $\mu(\varphi_k = a_{jk}) = b_{jk} > 0$, $b_{0k} + b_{1k} = 1$, $1 + a_{0k} a_{1k} = 0$, and $1 + a_{jk}^2 = 1/b_{jk}$, $j = 0, 1$; $k = 0, 1, \ldots$. In general, if f is integrable then

$$S_{2^n} f(\omega) = \int_{\Omega} f(t) \prod_{k=0}^{n-1} (1 + \varphi_k(\omega) \varphi_k(t))\, d\mu(t) .$$

Let $A = A(j_0, \ldots, j_{n-1}) = \{\omega \mid \varphi_k(\omega) = a_{j_k k}, \quad k = 0, \ldots, n-1\}$. Note that $\mu(A) = \prod_{k=0}^{n-1} (1 + a_{j_k k}^2)^{-1}$ using independence. Clearly, if $\omega \in A$ then by the above relations we have that

$$S_{2^n} f(\omega) = \int_A f(t)\, d\mu(t) / \mu(A) .$$

This establishes (8).

The convergence of $\{S_{2^n} f\}$ now follows from (8) and the martingale convergence theorem. The convergence is to $E(f \mid \varphi_0, \varphi_1, \ldots)$, which is f under the stated additional condition. The completeness of $\{\psi_n\}$ follows.

Theorem 8. *Suppose that $\varphi_0, \varphi_1, \ldots$ are identically distributed and that φ_0 has a two-point nonsymmetrical distribution ($\varphi_k \sim \varphi_0$, $k = 1, 2, \ldots$, but not $-\varphi_0 \sim \varphi_0$). Let $1 \leqq p < 2$. Then there is an $f \in L_p$ such that $\{S_n f\}$ diverges almost everywhere.*

If $\varphi_0, \varphi_1, \ldots$ are the Rademacher functions, then symmetry holds. For this case, Stein [*10*] has shown that there is a function f in L_1 such that $\{S_n f\}$ diverges almost everywhere. His approach does not seem to be applicable to the nonsymmetric case. Neither does a theorem of Alexits ([*1*], p. 250) apply here, since in the nonsymmetric case the sequence $\{\psi_n\}$ does not satisfy the second condition of his theorem.

Proof. Let $1 \leqq p \leqq 2$. Let $\mathscr{C}_p$ be the set of all sequences $(S_1 f, S_2 f, \ldots)$ such that $\int f = 0$ and $\int |f|^p \leqq 1$. Then $\mathscr{C}_p$ satisfies the condition C_p. For let $\{S_n f_k\} \in \mathscr{C}_p$, $k = 1, 2, \ldots$. Let $f_k' = T f_k$ where T is the conditional expectation $E(\cdot \mid \varphi_0, \varphi_1, \ldots)$. Then $(f_k', \psi_n) = (T f_k, \psi_n) = (f_k, T\psi_n) = (f_k, \psi_n)$, implying that

$\{S_n f_k'\} = \{S_n f_k\}$. There is a Baire function f_k'' such that $f_k' = f_k''(\varphi_0, \varphi_1, \ldots)$. Let n_{kj} be distinct positive integers satisfying $n_{k0} < n_{k1} < \cdots$. Let $g_k = f_k''(\varphi_{n_{k0}}, \varphi_{n_{k1}}, \ldots)$.
Then the sequences $(S_1 g_k, S_2 g_k, \ldots)$, $k = 1, 2, \ldots$, are independent and $S^* f_k \sim S^* g_k$, $k = 1, 2, \ldots$. Let $\{a_k\}$ be a real number sequence satisfying $\sum_{k=1}^{\infty} |a_k|^p = 1$. Then $\sum_{k=1}^{\infty} a_k S_n g_k$ converges almost everywhere, for all n, since the series consists of only a finite number of nonzero terms. By Lemma 2, stated below, we have that the series $\sum_{k=1}^{\infty} a_k g_k$ converges almost everywhere and in L_p norm to a function h satisfying

$$\int |h|^p \leqq 2^p \sum_{k=1}^{\infty} \int |a_k g_k|^p \leqq 2^p .$$

Since S_n is a bounded operator in L_p, it follows that $S_n h = \sum_{k=1}^{\infty} a_k S_n g_k$. Clearly, $\{S_n(h/2)\} \in \mathscr{C}_p$. This implies that $\mathscr{C}_p$ satisfies condition C_p.

Suppose that Theorem 8 is not true. Then there is a p_0 satisfying $1 \leqq p_0 < 2$ such that $\{S_n f\}$ converges on a set of positive measure if $f \in L_{p_0}$. That is, $\mathscr{C}_{p_0}$ satisfies the finiteness condition, and since $L_2 \subset L_{p_0}$ here, we also have that $\mathscr{C}_2$ satisfies the finiteness condition. Thus, by Theorem 2, there is a real number K such that if p is either p_0 or 2, we have that

$$\mu(S^* f > \lambda) \leqq K/\lambda^p, \quad \lambda > 0, \quad f \in \mathscr{C}_p .$$

From this we can easily deduce comparable inequalities for functions f in L_p not necessarily satisfying $\int f = 0$: We obtain the existence of a real number K such that if p is either p_0 or 2, then

$$\mu(S^* f > \lambda) \leqq K \int_{\Omega} |f|^p d\mu/\lambda^p$$

for $\lambda > 0$, $f \in L_p$. This implies by the MARCINKIEWICZ interpolation theorem ([12], II, page 112) that the norm $||S^*||_p$ of S^* is finite for $p_0 < p < 2$. Let $U_n f = (f, \psi_n)\psi_n$ for f in L_p. Then for $p_0 < p < 2$, we have that

$$\sup_n ||U_n||_p = \sup_n ||S_{n+1} - S_n||_p \leqq 2 \sup_n ||S_n||_p \leqq 2||S^*||_p < \infty .$$

Let $p_0 < p < 2$. Let $n = 2^{n_1} + \cdots + 2^{n_k}$, $n_1 > \cdots > n_k \geqq 0$. Then

$$||U_n||_p = ||\psi_n||_p \, ||\psi_n||_q = \prod_{j=1}^{k} ||\varphi_{n_j}||_p \, ||\varphi_{n_j}||_q = (||\varphi_0||_p \, ||\varphi_0||_q)^k ,$$

where $1/p + 1/q = 1$. But, by HÖLDER's inequality,

$$1 = ||\varphi_0||_2^2 < ||\varphi_0||_p \, ||\varphi_0||_q$$

since $|\varphi_0|$ has two nonzero values attained on sets of positive measure implying that $|\varphi_0|^p$ and $|\varphi_0|^q$ are not proportional. Thus, $\sup_n ||U_n||_p = \infty$, a contradiction, and the theorem is proved.

Lemma 2. *Let $1 \leqq p \leqq 2$. Suppose that $f_1, f_2, \ldots$ are independent functions on a probability space $(\Omega, \mathfrak{A}, \mu)$ such that $\sum_{k=1}^{\infty} \int |f_k|^p < \infty$ and $\int f_k = 0$, $k = 1, 2, \ldots$. Then the series $\sum_{k=1}^{\infty} f_k$ converges almost everywhere and in $L_p(\Omega, \mathfrak{A}, \mu)$ norm and*

$$\int \Big|\sum_{k=1}^{\infty} f_k\Big|^p \leqq 2^p \sum_{k=1}^{\infty} \int |f_k|^p .$$

Although this is well known, we sketch a proof. First suppose that $-f_k \sim f_k$ for all k. Then $\sum_{k=1}^{n} r_k(t) f_k \sim \sum_{k=1}^{n} f_k$ for each $t \in [0, 1]$ where $r_1, r_2, \ldots$ are the Rademacher functions (confer Section 1). Hence,

$$\begin{aligned}
(9) \qquad \int_\Omega \Big|\sum_{k=1}^{n} f_k\Big|^p d\mu &= \int_0^1 \int_\Omega \Big|\sum_{k=1}^{n} r_k(t) f_k(\omega)\Big|^p d\mu(\omega)\, dt \\
&\leqq \int_\Omega \left[\int_0^1 \Big|\sum_{k=1}^{n} r_k(t) f_k(\omega)\Big|^2 dt\right]^{p/2} d\mu(\omega) \\
&= \int_\Omega \left[\sum_{k=1}^{n} |f_k(\omega)|^2\right]^{p/2} d\mu(\omega) \\
&\leqq \int_\Omega \sum_{k=1}^{n} |f_k(\omega)|^p d\mu(\omega) .
\end{aligned}$$

For f_k nonsymmetric, one may suppose the existence of functions $g_1, g_2, \ldots$ (enlarge the space if necessary) such that $f_1, g_1, f_2, g_2, \ldots$ are independent and $f_k \sim g_k$ for all k. Since $\sum_{k=1}^{n} f_k - \sum_{k=1}^{n} g_k$ is mapped into $\sum_{k=1}^{n} f_k$ by the conditional expectation $E(\cdot \mid \sum_{k=1}^{n} f_k)$ and every conditional expectation has L_p norm 1, we have that

$$\int \Big|\sum_{k=1}^{n} f_k\Big|^p \leqq \int \Big|\sum_{k=1}^{n} f_k - \sum_{k=1}^{n} g_k\Big|^p$$

which is less than or equal to $2^p \sum_{k=1}^{n} \int |f_k|^p$ by (9) applied to $\{f_k - g_k\}$ and an elementary inequality. The martingale convergence theorem now implies the desired result.

If $\{\varphi_k\}$ does not satisfy the two-point condition, then $\{\psi_n\}$ is not complete. For example, if φ_0 is not two-point, then there is a nonzero function f, orthogonal to every ψ_n, of the form $f(\omega) = 1, c_1, c_2$ according as $\varphi_0(\omega)$ is less than a_1, is in $[a_1, a_2]$, is greater than a_2, respectively.

As a further contrast to the results of Theorem 7 (and of Corollary 1) we have the following.

Theorem 9. *Let $1 \leqq p < 2$, $1/p + 1/q = 1$. Suppose that $\varphi_0, \varphi_1, \ldots$ are identically distributed, $\varphi_0 \in L_q$, and that φ_0 does not have a two-point distribution. Then there is an $f \in L_p$ such that $\{S_{2n} f\}$ diverges almost everywhere.*

The proof of this theorem will be omitted since it is virtually the same as that of Theorem 8. The greatest change is in using the fact that here $\sup_n \| S_{2^n} \|_r = \infty$ for $p < r < 2$. To see this, let $1/r + 1/s = 1$, $g_k = (\varphi_{2k} - \varphi_{2k+1})/\sqrt{2}$, $f_k = (\operatorname{sign} g_k)(|g_k|/\|g_k\|_s)^{s-1}$, and $f = \prod_{k=0}^{n-1} f_k$. Then $\|f\|_r = 1$ and

$$S_{2^{2n}} f(\omega) = \prod_{k=0}^{n-1} \int_\Omega f_k(t)\,(1 + \varphi_{2k}(\omega)\,\varphi_{2k}(t))\,(1 + \varphi_{2k+1}(\omega)\,\varphi_{2k+1}(t))\,d\mu(t)$$

$$= \prod_{k=0}^{n-1} g_k(\omega) \int f_k(t)\, g_k(t)\, d\mu(t);$$

hence,

$$\| S_{2^{2n}} \|_r \geqq \| S_{2^{2n}} f \|_r = (\|g_0\|_r \, \|g_0\|_s)^n .$$

Since $|g_0|$ is nonconstant on the set where $g_0 \neq 0$, we have that $\|g_0\|_r \, \|g_0\|_s > \|g_0\|_2^2 = 1$, and the desired result is implied.

Added in proof, May 9, 1964: S. SAWYER has recently discovered still another condition assuring that almost everywhere convergence will imply a maximal inequality. This condition seems to be applicable in certain ergodic theory contexts rather different from those discussed here.

References

[1] ALEXITS, G.: Convergence Problems of Orthogonal Series. New York: Pergamon 1961.
[2] BURKHOLDER, D. L.: Successive conditional expectations of an integrable function. Ann. math. Statistics **33**, 887—893 (1962).
[3] —, and Y. S. CHOW: Iterates of conditional expectation operators. Proc. Amer. math. Soc. **12**, 490—495 (1961).
[4] DOOB, J. L.: Stochastic Processes. New York: Wiley 1953.
[5] DUNFORD, N., and J. T. SCHWARTZ: Convergence almost everywhere of operator averages. J. Rat. Mech. Analysis **5**, 129—178 (1956).
[6] — — Linear Operators, I. New York: Interscience 1958.
[7] HOPF, E.: The general temporally discrete Markoff process. J. Rat. Mech. Analysis **3**, 13—45 (1954).
[8] ROTA, G.-C.: An "alternierendes Verfahren" for general positive operators. Bull. Amer. Math. Soc. **68**, 95—102 (1962).
[9] SCHWARTZ, J. T.: Another proof of E. Hopf's ergodic lemma. Commun. pure appl. Math. **12**, 399—401 (1959).
[10] STEIN, E. M.: On limits of sequences of operators. Ann. of Math. **74**, 140—170 (1961).
[11] — On the maximal ergodic theorem. Proc. nat. Acad. Sci. USA. **47**, 1894—1897 (1961).
[12] ZYGMUND, A.: Trigonometric Series, I, II. Cambridge 1959.

Department of Mathematics
University of Illinois
Urbana, Illinois 61803

(Received August 20, 1963)