The Annals of Probability
1987, Vol. 15, No. 4, 1569–1585

A TWO-PARAMETER MAXIMAL ERGODIC THEOREM WITH DEPENDENCE

BY TERRY R. MCCONNELL

Syracuse University

Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be independent sequences of i.i.d. $U(0,1)$ random variables. We characterize completely those Borel functions F on $[0,1]^2$ for which the strong law of large numbers and the maximal ergodic theorem hold for the doubly indexed family $(1/nm)\Sigma_{i \le n, j \le m} F(X_i, Y_j)$.

1. Introduction. Let $Z, Z_1, Z_2, \dots$ be i.i.d. random variables and denote by S_n the nth partial sum of the Z_i. Then $\sup_n(1/n)|S_n|$ is finite almost surely or, equivalently, the strong law of large numbers holds, if and only if $|Z|_1 = E|Z| < \infty$. It is also well known that the stronger condition $E(\sup_n(1/n)|S_n|) < \infty$ holds if and only if Z belongs to the class $L \log_+ L$. This sharp analog of Hopf's maximal ergodic theorem in the case of i.i.d. summands is due to Marcinkiewicz and Zygmund [9] and Burkholder [2]. More recently, several authors have obtained generalizations of these results to the case of multiply indexed families (see, e.g., [6] and [15]) but under the assumption that the summands are mutually independent. The primary goal of the present paper is to prove a strong law of large numbers and a maximal ergodic theorem for a doubly indexed family exhibiting a nontrivial pattern of dependence.

Recall that the class $L \log_+ L$ consists of those random variables Z for which the norm $|Z|_{L \log_+ L}$ defined by

$$|Z|_{L \log_+ L} = E|Z|\left(1 + \log_+ \frac{|Z|}{|Z|_1}\right) \tag{1.1}$$

is finite. (Actually $|\ |_{L \log_+ L}$ is not a norm since it fails to satisfy the triangle inequality. However, it is easily seen to be comparable in size with any of the standard Orlicz-type norms on $L \log_+ L$.) For functions F on $[0,1]^2$ we denote by $\|F\|_{L \log_+ L}$ the iterated $L \log_+ L$ norm, i.e., $\|F\|_{L \log_+ L} = |f|_{L \log_+ L}$, where $f(x) = |F(x, \cdot)|_{L \log_+ L} + |F(\cdot, x)|_{L \log_+ L}$. We shall also need several other norm-like quantities associated with functions of two variables. These are defined by

$$\delta(F) = \int_{[0,1]^2} |F(x,y)|\left(1 + \log_+\left(\frac{|F|_1 |F(x,y)|}{|F(x,\cdot)|_1 |F(\cdot,y)|_1}\right)\right) dx\, dy,$$

$$\Delta(F) = \int_{[0,1]^2} |F(x,y)| \log_+^2\left(\frac{|F|_1 |F(x,y)|}{|F(x,\cdot)|_1 |F(\cdot,y)|_1}\right) dx\, dy$$

and

$$\|F\| = \|F\|_{L \log_+ L} + \Delta(F).$$

Received September 1985; revised April 1986.
AMS 1980 *subject classifications*. Primary 28D05; secondary 60G60.
Key words and phrases. Maximal ergodic theorem, strong law of large numbers, two-parameter martingales, decoupling.

Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be independent sequences of i.i.d. $U(0,1)$ random variables and let $\varepsilon_1, \varepsilon_2, \ldots$ and $\varepsilon'_1, \varepsilon'_2, \ldots$ be independent Rademacher sequences which are also independent of the X_i and Y_i. For each Borel function F on $[0,1]^2$ define

$$^*F^* = \sup_{i,j} \frac{1}{ij}\left|F(X_i, Y_j)\right|$$

and

$$^*S^*(F) = \sup_{n,m} \frac{1}{nm}\left|\sum_{i=1}^{n}\sum_{j=1}^{m} F(X_i, Y_j)\right|.$$

THEOREM 1. *The following statements concerning the Borel function F are equivalent*:

(1.2) $$^*F^* < \infty \quad a.s.,$$

(1.3) $$^*S^*(F) < \infty \quad a.s.,$$

(1.4) $$\sum_{i=1}^{n}\sum_{j=1}^{m} \frac{\varepsilon_i \varepsilon'_j}{ij} F(X_i, Y_j) \quad \text{converges a.s. as } n \wedge m \to \infty,$$

(1.5) $$\delta(F) < \infty$$

and

(1.6) $$\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} F(X_i, Y_j) \to \int_{[0,1]^2} F(x, y)\, dx\, dy \quad a.s. \text{ as } n \wedge m \to \infty.$$

Note that the equivalence (1.5) $\Leftrightarrow$ (1.6) is the strong law of large numbers mentioned previously. In the rest of this paper, convergence of a double series means convergence in the sense of (1.4).

In our second result, the notation $A \approx B$ means that there are absolute constants c_1 and c_2 such that $c_1 A \le B \le c_2 A$.

THEOREM 2. *Let F be a Borel function on* $[0,1]^2$. *Then we have*

$$E^*F^* \approx E^*S^*(F) \approx E\left|\sum_{i,j} \frac{\varepsilon_i \varepsilon'_j}{ij} F(X_i, Y_j)\right|.$$

Moreover, all three quantities are comparable with $\|F\|$.

We shall prove Theorems 1 and 2 in Section 3. In Section 2, we collect some useful background information and in Section 4, we pose some open problems.

2. Background results. The purpose of this section is to present some preliminary results from classical probability theory in a convenient form and to discuss some more recent results on double random series and multiple stochastic integration.

Recall that $\{X_i\}$ and $\{Y_i\}$ denote independent sequences of i.i.d. $U(0,1)$ random variables and that $\{\varepsilon_i\}$ and $\{\varepsilon_i'\}$ denote independent Rademacher sequences which are independent of the X_i and Y_i. For Borel functions f on $[0,1]$ define

$$f^* = \sup_i |f(X_i)|/i$$

and

$$s^*(f) = \sup_n \frac{1}{n}\left|\sum_{i=1}^n f(X_i)\right|.$$

PROPOSITION 2.1. *The statements* (2.1)–(2.4) *concerning a Borel function f are equivalent*:

(2.1) $$f^* < \infty \quad a.s.,$$

(2.2) $$s^*(f) < \infty \quad a.s.,$$

(2.3) $$\sum_{i=1}^{\infty} \varepsilon_i \frac{f(X_i)}{i} \quad \text{converges } a.s.$$

and

(2.4) $$f \in L^1[0,1].$$

Moreover, the following equivalences hold:

(2.5) $$Ef^* \approx Es^*(f) \approx E\left|\sum_{i=1}^{\infty} \varepsilon_i \frac{f(X_i)}{i}\right| \approx |f|_{L \log_+ L}.$$

The last equivalence in (2.5) is contained in Proposition 4.2 of [3]. The remaining parts of (2.5) are well known.

In the course of the proof of Theorems 1 and 2 we shall be concerned with convergence of sums of the form $\sum_{i,j}(\varepsilon_i\varepsilon_j'/ij)F(X_i, Y_j)$ or, more generally, with convergence of sums of the form $\sum_{i,j} a_{ij}\varepsilon_i\varepsilon_j'$, $a_{ij} \in \mathbb{R}$. The next proposition summarizes the properties of double Rademacher series needed later. In particular, it yields a "contraction principle" and also shows that double Rademacher series share with single series the property that L^p convergence, convergence almost surely and convergence in probability are all equivalent.

PROPOSITION 2.2. *The double series $\sum_{i,j} a_{ij}\varepsilon_i\varepsilon_j'$ converges almost surely if and only if $\sum_{i,j} a_{ij}^2 < \infty$. Moreover, we have the following extension of Khintchine's inequalities*:

(2.6) $$E\left|\sum_{i,j} a_{ij}\varepsilon_i\varepsilon_j'\right|^p \approx \left(\sum_{i,j} a_{ij}^2\right)^{p/2}, \qquad 0 < p < \infty.$$

This result is implicitly contained in a paper of Bonami [1]. More recent and, perhaps, more convenient references are [7, Theorem 2.5] for the first statement and [11, Proposition 2.2] for (2.6).

Several parts of the proofs of Theorems 1 and 2 rely on recent results concerning double integration with respect to symmetric stable processes, and the remainder of this section is devoted to that subject. Let X_t and Y_t be independent copies of the standard symmetric stable process of index α in $(0,2)$, i.e., X_t has stationary, independent increments and

$$Ee^{i\lambda X_t} = e^{-t|\lambda|^\alpha}, \qquad \lambda \in \mathbb{R}.$$

We may assume also that the sample paths of X_t are right continuous with left limits.

If F is a step function on $[0,1]^2$ we define the double integral

$$I_\alpha(F) = \int_0^1 \int_0^1 F(t,s)\, dX_t\, dY_s$$

in the obvious way. What is the widest class of integrands F for which I_α is well-defined? Of course the answer to this question depends on *how* the integral is defined, but it turns out that all reasonable definitions of I_α and of the more natural integral

$$\int_0^1 \int_0^1 F(t,s) I_{\{t \neq s\}}\, dX_t\, dX_s$$

lead to the same condition on F. More precisely, we have

PROPOSITION 2.3. *Let α satisfy $1 \le \alpha < 2$. The following conditions on a Borel function F on $[0,1]^2$ are equivalent:*

(2.7) *There exist step functions F_n converging to F in measure for which $I_\alpha(F_n)$ converges in probability.*

(2.8) *The processes $s \mapsto \int_0^1 F(t,s)\, dX_t$ and $t \mapsto \int_0^1 F(t,s)\, dY_s$ have versions with sample paths in $L^\alpha[0,1]$.*

(2.9) $$\sum_{i,j} \frac{\varepsilon_i \varepsilon_j'}{(ij)^{1/\alpha}} F(X_i, Y_j) \quad \textit{converges a.s. (as in (1.4)).}$$

(2.10) $$\delta_\alpha(F) = \int_0^1 \int_0^1 |F(t,s)|^\alpha \left(1 + \log_+ \frac{|F|_\alpha^\alpha |F(t,s)|^\alpha}{\int_0^1 |F(t,s')|^\alpha\, ds' \int_0^1 |F(t',s)|^\alpha\, dt'}\right) dt\, ds < \infty.$$

Moreover, if $\delta_\alpha(F) < \infty$ the step functions in (2.7) *may be chosen so that $\delta_\alpha(F - F_n) \to 0$.*

The reason for the restriction on α is that the proof relies on the spectral theory of stable probability measures on Banach spaces, and $L^\alpha[0,1]$ is not a Banach space when $0 < \alpha < 1$. The equivalence of (2.8) and (2.10) is proved in [14, Theorem 6.2]. A proof of the equivalence of (2.7), (2.10) and the last statement in the proposition may be obtained by combining the result of [10] with [12, Theorem 7.3]. We now sketch a proof of the equivalence (2.8) $\Leftrightarrow$ (2.9) in the case $\alpha = 1$ of interest here. The proof is essentially contained in [12] and is based on a novel representation of Y_t due to LePage. Let $e_1, e_2, \ldots$ be i.i.d.

standard exponential random variables which are independent of the $U(0,1)$ random variables Y_j previously introduced and also independent of the Rademacher sequence ε_j'. Put $\Gamma_j = e_1 + \cdots + e_j$. Then the series [8]

$$\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{\Gamma_j} 1_{[0, Y_j)}(s)$$

converges a.s. and defines a process stochastically equivalent to Y_s. Now, by [12, Proposition 3.3] the first statement in (2.8) is equivalent to

$$\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{\Gamma_j} F(t, Y_j) \quad \text{converges in } L^1[0,1] \quad \text{a.s.}$$

This is equivalent, by Kahane's contraction principle, to the almost sure convergence in $L^1[0,1]$ of

$$\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} F(\cdot, Y_j). \tag{2.11}$$

By Proposition 2.1 [(2.4) ⇒ (2.3)] we have that (2.11) implies almost sure convergence of

$$\sum_{i=1}^{\infty} \frac{\varepsilon_i}{i} \sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} F(X_i, Y_j).$$

Now choose $n < m$ and define Z_i and W_i by

$$Z_i + W_i = \sum_{j=1}^{\infty} \frac{\varepsilon_j'}{ij} F(X_i, Y_j)$$

and

$$Z_i = \sum_{j=n}^{m} \frac{\varepsilon_j'}{ij} F(X_i, Y_j).$$

Then since Z_i and W_i are symmetric and conditionally independent given the X_k and Y_k, we have for every $N < M$ and $\varepsilon > 0$,

$$P\left(\left|\sum_{i=N}^{M} \varepsilon_i Z_i\right| > \varepsilon\right) = 2P\left(\sum_{i=N}^{M} \varepsilon_i Z_i > \varepsilon\right) \le 4P\left(\sum_{i=N}^{M} \varepsilon_i Z_i > \varepsilon,\ \sum_{i=N}^{M} \varepsilon_i W_i \ge 0\right)$$
$$\le 4P\left(\left|\sum_{i=N}^{M} \varepsilon_i (Z_i + W_i)\right| > \varepsilon\right).$$

This and a similar argument with the roles of i and j reversed show that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\varepsilon_i \varepsilon_j'}{ij} F(X_i, Y_j)$$

converges in probability as $n \wedge m \to \infty$. Choosing a subsequence (n_i, m_i) with $n_i \wedge m_i \to \infty$ along which we have almost sure convergence and applying

Proposition 2.2 show that

$$\sum_{i,j} \frac{F^2(X_i, Y_j)}{(ij)^2} < \infty \quad \text{a.s.}$$

Thus (2.9) holds by another appeal to Proposition 2.2.

For the converse implication, (2.9) and Proposition 4.1 of [3] yield

$$\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} F(x, Y_j) \quad \text{converges a.s. for a.e. } x,$$

(2.12)

$$\sum_{i=1}^{\infty} \frac{\varepsilon_i}{i} \left(\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} F(X_i, Y_j) \right) \quad \text{converges a.s.}$$

Statement (2.11) and, hence, the first statement of (2.8) now follow from Proposition 2.1 [(2.3) ⇒ (2.4)] and the Itô–Nisio theorem. The second statement in (2.8) follows similarly by interchanging the roles of x and y.

REMARK 2.1. The functional δ_α already defined is not a norm since it fails to satisfy the triangle inequality. However, it is comparable in size to a p-homogeneous quasinorm for any $0 < p < \alpha$ [12, Section 7] and the class of functions defined by any of the equivalent conditions in Proposition 2.3 forms a Banach space. Furthermore, we have $\delta_\alpha(F) < \infty$ implies $|F|_\alpha < \infty$.

REMARK 2.2. It would be interesting to have a direct proof of the equivalence (2.9) and (2.10) which avoids the introduction of stable processes.

3. Proofs of the main theorems. In this section we present proofs of Theorems 1 and 2.

PROOF OF THEOREM 1. Each of the conditions (1.2)–(1.6) entails the condition $F \in L^1([0,1]^2)$. Therefore, the one-parameter results (i.e., Proposition 2.1) may be used to control the behavior of $|F(X_i, Y_j)|/ij$ on any finite number of horizontal or vertical lines in the parameter space. With the aid of this observation, the implications (1.3) ⇒ (1.2), (1.4) ⇒ (1.2) and (1.6) ⇒ (1.3) follow easily. The equivalence (1.4) ⇔ (1.5) has been proved in Proposition 2.3. The remainder of the argument will be presented in the following order: (1.2) ⇒ (1.3), (1.2) ⇒ (1.4) and (1.5) ⇒ (1.6).

(1.2) ⇒ (1.3). For a measurable function F on $[0,1]^2$, define on $\Omega_X \times [0,1]$,

$$^*F(y) = {^*F}(\omega, y) = \sup_i \frac{|F(X_i(\omega), y)|}{i},$$

and on $[0,1] \times \Omega_Y$,

$$F^*(x) = F^*(x, \omega) = \sup_j \frac{|F(x, Y_j(\omega))|}{j}.$$

The equivalence of statements (2.1), (2.2) and (2.4) in Proposition 2.1 may be used together with Fubini's theorem to obtain the following string of implications:

$$\text{“}{}^*F^* < \infty \text{ a.s.”} \Rightarrow \text{“}\int_0^1 F^*(x,\omega)\,dx < \infty \text{ a.s.”} \Rightarrow \text{“}s^*(F^*) < \infty \text{ a.s.”}$$

But $s^*(F^*) \geq \sup_j (1/j) s^*(F(\cdot, Y_j))$. The proof of (1.2) ⇒ (1.3) is completed by the next string of implications which, again, follow from Proposition 2.1:

$$\text{“}\sup_j \frac{1}{j} s^*\big(F(\cdot, Y_j)\big) < \infty \text{ a.s.”} \Rightarrow \text{“}\int_0^1 s^*(F(\cdot, y))\,dy < \infty \text{ a.s.”}$$

$$\Rightarrow \text{“}{}^*S^*(F) < \infty \text{ a.s.”}$$

(1.2) ⇒ (1.4). As in the preceding proof of (1.2) ⇒ (1.3), we conclude that $\int_0^1 F^*(x,\omega)\,dx < \infty$, a.s. This, together with Fubini's theorem and Proposition 2.1, imply that $\sum_{i=1}^\infty (\varepsilon_i/i) F^*(X_i, \omega)$ converges, a.s. It follows that

$$\sum_{i=1}^{\infty} \frac{\left(F^*(X_i,\omega)\right)^2}{i^2} < \infty \quad \text{a.s.};$$

hence, that

$$\sup_j \frac{1}{j} \left(\sum_{i=1}^{\infty} \frac{F^2(X_i, Y_j)}{i^2} \right)^{1/2} < \infty \quad \text{a.s.}$$

By Proposition 2.1, (2.1) ⇒ (2.4) ⇒ (2.3), we conclude that

$$\sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} \left(\sum_{i=1}^{\infty} \frac{F^2(X_i, Y_j)}{i^2} \right)^{1/2} \quad \text{converges a.s.}$$

and, hence,

$$\sum_{i,j} \frac{F^2(X_i, Y_j)}{(ij)^2} < \infty \quad \text{a.s.}$$

Finally, this implies (1.4) by Proposition 2.2.

(1.5) ⇒ (1.6). We first show that $\delta(G_n) \to 0$ implies ${}^*S^*(G_n) \to_P 0$. If this were false, we could find nonnegative functions G_n and $\varepsilon > 0$ so that $P({}^*S^*(G_n) > n) > \varepsilon$ and $\delta(G_n) \leq 2^{-n}$. By Remark 2.1 the series $\sum_{n=1}^\infty G_n$ converges a.e. to a function G satisfying $\delta(G) < \infty$. But $P({}^*S^*(G) = +\infty) \geq P({}^*S^*(G_n) > n, \text{ i.o.}) > 0$, and this contradicts the implication (1.5) ⇒ (1.3) already proved.

Now suppose $\delta(F) < \infty$ and choose step functions F_n such that $\delta(F - F_N) \to 0$ as in Proposition 2.3. For the F_N, we clearly have that

$$\frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} F_N(X_i, Y_j)$$

converges almost surely as $n \wedge m \to \infty$ to

$$\int_{[0,1]^2} F_N(x, y)\, dx\, dy.$$

Since ${}^*S^*(F - F_N) \to_P 0$ and $|F_N - F|_1 \to 0$, the desired conclusion follows. □

PROOF OF THEOREM 2. The inequality $E^*F^* \le E^*S^*(F)$ is clear. To obtain the reverse inequality, we have by (2.5),

$$\begin{aligned} E^*F^* &\approx E|{}^*F(\cdot)|_{L \log_+ L} \approx E \sup_N \frac{1}{N} \sum_{j=1}^{N} {}^*F(Y_j) \\ &\ge E \sup_i \frac{1}{i} \sup_N \frac{1}{N} \left| \sum_{j=1}^{N} F(X_i, Y_j) \right| \\ &= E\left(E \sup_i \frac{1}{i} \sup_N \frac{1}{N} \left| \sum_{j=1}^{N} F(X_i, Y_j) \right| \Bigg| (Y_j) \right) \ge cE^*S^*(F). \end{aligned}$$

Next we prove the equivalence

$$E\left| \sum_{i,j} \frac{\varepsilon_i \varepsilon_j'}{ij} F(X_i, Y_j) \right| \approx E^*F^*. \tag{3.1}$$

By (2.5), Khintchine's inequality and (2.6)

$$\begin{aligned} E^*F^* &\approx E\left| \sum_{j=1}^{\infty} \frac{\varepsilon_j'}{j} {}^*F(Y_j) \right| \\ &\approx E\left(\sum_{j=1}^{\infty} \frac{({}^*F(Y_j))^2}{j^2} \right)^{1/2} \ge E \sup_i \frac{1}{i} \left(\sum_{j=1}^{\infty} \frac{(F(X_i, Y_j)^2}{j^2} \right)^{1/2} \\ &\approx E\left| \sum_{i=1}^{\infty} \frac{\varepsilon_i}{i} \left(\sum_{j=1}^{\infty} \frac{F(X_i, Y_j)^2}{j^2} \right)^{1/2} \right| \\ &\approx E\left| \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{F(X_i, Y_j)^2}{(ij)^2} \right|^{1/2} \approx E\left| \sum_{i,j} \frac{\varepsilon_i \varepsilon_j'}{ij} F(X_i, Y_j) \right|. \end{aligned}$$

The reverse inequality follows from (2.6):

$$E\left| \sum_{i,j} \frac{\varepsilon_i \varepsilon_j'}{ij} F(X_i, Y_j) \right| \approx E\left(\sum_{i,j} \frac{F(X_i, Y_j)^2}{(ij)^2} \right)^{1/2} \ge E^*F^*.$$

The equivalence of E^*F^* and $\|F\|$ is much more difficult. It will be established in the following steps.

STEP 1. Put

$$|F|_* = \sum_{j=1}^{\infty} \int_0^1 E\left\{ \frac{|F(X_1, y)|}{j} \log_+ \left(\frac{|F(X_1, y)|}{2(|F(X_1, \cdot)|_1 \vee \beta j)} \right); |F(X_1, y)| > \lambda(y) j \right\} dy,$$

where $\lambda(y) = |F(\cdot, y)|_{L \log_+ L}$ and $\beta = \|F\|_{L \log_+ L}$.

We show that there is a constant C_1 such that

$$E^*F^* \le C_1\big(\|F\|_{L \log_+ L} + |F|_*\big).$$

STEP 2. We show that there is a constant C_2 such that

$$\|F\|_{L \log_+ L} + |F|_* \le C_2 E^*F^*.$$

STEP 3. We prove that

$$\|F\|_{L \log_+ L} + |F|_* \approx \|F\|.$$

We shall use the following elementary inequalities repeatedly: For $A > 0$ and $B > 0$,

$$\log_+(A \vee B) \le \log_+ A + \log_+ B, \quad \log_+(AB) \le \log_+ A + \log_+ B, \tag{3.2}$$

$$A \log_+(B/A) \le B/e, \qquad A \log_+^2(B/A) \le B/e^2, \tag{3.3}$$

and

$$A \log_+ B \le A \log_+ A + B/e, \qquad e = 2.718\dots. \tag{3.4}$$

STEP 1 (discussion). By (2.5) we have that

$$E^*F^* \approx \int_0^1 E^*F(y)\left(1 + \log_+ \frac{{}^*F(y)}{|{}^*F|_1}\right) dy. \tag{3.5}$$

Now

$$\int_0^1 E^*F(y)\, dy \approx \int_0^1 \lambda(y)\, dy \le \beta$$

by (2.5). Thus, it is sufficient to show that

$$\int_0^1 E^*F(y) \log_+ \frac{{}^*F(y)}{|{}^*F|_1}\, dy \le C(\beta + |F|_*).$$

Now, by (3.2),

$$\int_0^1 E^*F(y) \log_+ \frac{{}^*F(y)}{|{}^*F|_1}\, dy \le \int_0^1 E^*F(y) \log_+ \frac{{}^*F(y)}{2(|{}^*F|_1 \vee \beta)}\, dy + \int_0^1 E^*F(y) \log_+\left(\frac{2(|{}^*F|_1 \vee \beta)}{|{}^*F|_1}\right) dy.$$

The second term becomes [see (3.3)]

$$E|{}^*F|_1 \log_+ \left(\frac{2(|{}^*F|_1 \vee \beta)}{|{}^*F|_1} \right) \le \left(\frac{4}{e} \right) \beta.$$

Thus, we need only prove that

$$\int_0^1 E{}^*F(y) \log_+ \frac{{}^*F(y)}{2(|{}^*F|_1 \vee \beta)}\, dy \le c(|F|_* + \beta). \tag{3.6}$$

(Note that the denominator of the argument of $\log_+$ is random.)

By (3.4),

$$\int_0^1 \lambda(y) E \log_+ \frac{{}^*F(y)}{|{}^*F|_1}\, dy \le \int_0^1 \lambda(y) \log_+ \frac{\lambda(y)}{\beta}\, dy + \frac{\beta}{e} \int_0^1 E\left(\frac{{}^*F(y)}{|{}^*F|_1} \right) dy;$$

hence

$$\int_0^1 \lambda(y) E \log_+ \frac{{}^*F(y)}{|{}^*F|_1}\, dy \le \left(1 + \frac{1}{e} \right) \beta. \tag{3.7}$$

The remainder of the proof is based on a simple pointwise upper bound which will also be used in Step 2: For $g \ge 0$ Borel and any constant $\lambda > 0$, we have

$$\sup_i \frac{g(X_i)}{i} \le \lambda + \sum_{i=1}^{\infty} \frac{g(X_i)}{i} I(g(X_i) > \lambda i).$$

This estimate is actually quite sharp since integration of both sides and appropriate choice of λ leads to one of the equivalences of (2.5) after standard manipulations.

This estimate, applied for each y with $\lambda = \lambda(y)$ to the first appearance of ${}^*F$ in (3.6), together with (3.7) show that it is enough to prove

$$\sum_{i=1}^{\infty} \int_0^1 E\left[\frac{|F(X_i, y)|}{i} \log_+ \frac{{}^*F(y)}{2(|{}^*F|_1 \vee \beta)};\ |F(X_i, y)| > \lambda(y) i \right] dy \le c(\beta + |F|_*). \tag{3.8}$$

Put

$$H_i(y) = \sup_{j \ne i} \frac{|F(X_j, y)|}{j}.$$

Since H_i is independent of X_i, we obtain from (3.2) that the expression in (3.8) is dominated by

$$\sum_{i=1}^{\infty} \int_0^1 E\left(\frac{|F(X_i, y)|}{i};\ |F(X_i, y)| > i\lambda(y) \right) \left(E \log_+ \frac{H_i(y)}{2|H_i|_1} \right) dy$$
$$+ \sum_{i=1}^{\infty} \int_0^1 E\left(\frac{|F(X_i, y)|}{i} \log_+ \left(\frac{|F(X_i, y)|}{2(|F(X_i, \cdot)|_1 \vee i\beta)} \right);\ |F(X_i, y)| > i\lambda(y) \right) dy$$
$$= (\mathrm{I}) + (\mathrm{II}),$$

say. The second term reduces to $|F|_*$. To estimate (I), note that by relabeling indices we have an estimate,

$$E\log_+\frac{H_i(y)}{2|H_i|_1}\le E\log_+\frac{2^*F(y)}{|^*F|_1}:=M(y),$$

which is independent of i. Thus, by a standard calculation,

$$\begin{aligned}(\mathrm{I}) &\le \int_0^1 M(y)\sum_{i=1}^{\infty}E\left(\frac{|F(X_1,y)|}{i};\,|F(X_1,y)|>i\lambda(y)\right)dy\\ &\le \int_0^1 M(y)E|F(X_1,y)|\log_+\frac{|F(X_1,y)|}{\lambda(y)}\,dy\\ &\le \int_0^1 M(y)E|F(X_1,y)|\log_+\frac{|F(X_1,y)|}{|F(\cdot,y)|_1}\,dy\\ &\le \int_0^1 M(y)\lambda(y)\,dy.\end{aligned}$$

We conclude from (3.7) that (I) $\le c\beta$, concluding Step 1.

STEP 2 (discussion). We may assume $E^*F^* < \infty$. Then, by (2.5), we have

$$\begin{aligned}E^*F^* &\ge E\left(\sup_i\frac{1}{i}E\left(\sup_j\frac{1}{j}|F(X_i,Y_j)|\,\Big|\,(X_k)_{k=1}^{\infty}\right)\right)\\ &\ge cE\sup_i\frac{1}{i}|F(X_i,\cdot)|_{L\log_+L}.\end{aligned}\tag{3.9}$$

Reversing the roles of the variables we obtain

$$E^*F^*\ge\frac{C}{2}\|F\|_{L\log_+L}=\frac{C}{2}\beta.$$

Now define for each y two sequences of events,

$$\{A_j(y)\}_{j=0}^{\infty}\quad\text{and}\quad\{B_j(y)\}_{j=1}^{\infty},$$

by $A_0(y)=\Omega$, and

$$A_j(y)=\left\{\sup_{i\le j}\frac{1}{i}|F(X_i,y)|\le\nu\lambda(y)\right\},$$

$$B_j(y)=A_{j-1}\cap\left\{\left|F(X_j,y)\right|>\nu\lambda(y)j\right\},\qquad j=1,2,\dots,$$

where an integer ν is chosen so that $E^*F(y)\le(\nu/4)\lambda(y)$. Note that the B_j are pairwise disjoint and that Chebyshev's inequality gives

$$P(A_j)\ge 3/4.\tag{3.10}$$

Now, by (3.5),

$$E^*F^* \geq c\int_0^1 E^*F(y)\log_+\left(\frac{{}^*F(y)}{|{}^*F|_1}\right)dy$$

$$\geq c\sum_{j=1}^{\infty}\int_0^1 E\left\{\frac{|F(X_j, y)|}{j}\log_+\left(\frac{|F(X_j, y)|/j}{|{}^*F|_1}\right); B_j\right\}dy$$

$$\geq c\sum_{j=1}^{\infty}\int_0^1 E\left\{\frac{|F(X_j, y)|}{j}\log_+\left(\frac{|F(X_j, y)|/j}{2\left(|F(X_j, \cdot)|_1/j \vee |H_j|_1\right)}\right); B_j\right\}dy,$$

where we have used the elementary inequality

$$|f \vee g|_1 \leq 2(|f|_1 \vee |g|_1)$$

for nonnegative functions f and g in the last step.

By the preceding choice of the integer ν, we have that $E|{}^*F|_1 \leq \nu\beta/4$. Then, by (3.10),

$$P(|H_j|_1 \leq \nu\beta, A_{j-1}) \geq P(|H_j|_1 \leq \nu\beta) + P(A_{j-1}) - 1 \geq \tfrac{1}{2}.$$

Hence, the previous sum dominates:

$$\frac{1}{2}\sum_{j=1}^{\infty}\int_0^1 E\left\{\frac{|F(X_j, y)|}{j}\log_+\frac{|F(X_j, y)|}{2\left(|F(X_j, \cdot)|_1 \vee \nu j\beta\right)}; |F(X_j, y)| > \nu j\lambda(y)\right\}dy$$

$$= \frac{1}{2}\sum_{j=1}^{\infty}\int_0^1 E\left\{\frac{|F(X_{\nu j}, y)|}{j}\log_+\frac{|F(X_{\nu j}, y)|}{2\left(|F(X_{\nu j}, \cdot)|_1 \vee \nu j\beta\right)}; |F(X_{\nu j}, y)| > \nu j\lambda(y)\right\}dy$$

$$\geq \frac{1}{2\nu}\sum_{j=\nu}^{\infty}\int_0^1 E\left\{\frac{|F(X_j, y)|}{j}\log_+\frac{|F(X_j, y)|}{2\left(|F(X_j, \cdot)|_1 \vee j\beta\right)}; |F(X_j, y)| > j\lambda(y)\right\}dy$$

$$= \frac{1}{2\nu}\sum_{j=\nu}^{\infty}\int_0^1 E\left\{\frac{|F(X_1, y)|}{j}\log_+\frac{|F(X_1, y)|}{2\left(|F(X_1, \cdot)|_1 \vee j\beta\right)}; |F(X_1, y)| > j\lambda(y)\right\}dy.$$

This last expression is exactly $1/2\nu|F|_*$ with the first $\nu - 1$ terms missing. It is easy to check that each of the missing terms is dominated by β. Thus, $|F|_* \le 2\nu c E({}^*F^*) + (\nu - 1)\beta$. We conclude from (3.9) that

$$|F|_* + \beta \le (2\nu c + \nu) E({}^*F^*)$$

and Step 2 is complete.

STEP 3 (discussion). We shall write $A \sim B$ to denote equivalence up to terms of order β, i.e., a relationship of the form

$$c_0 B - c_1\beta \le A \le c_2 B + c_3\beta \quad \text{for some } c_i \ge 0.$$

It is then enough to prove

$$|F|_* = \sum_{j=1}^{\infty} \int_0^1 E\left\{ \frac{|F(X_1, y)|}{j} \log_+\left(\frac{|F(X_1, y)|}{2\big(|F(X_1, \cdot)|_1 \vee j\beta\big)} \right); \right.$$

$$\left. |F(X_1, y)| \ge j\lambda(y) \right\} dy \tag{3.11}$$

$$\sim \Delta(F) = \int_0^1 \int_0^1 |F(x, y)| \log_+^2 \left\{ \frac{|F|_1 |F(x, y)|}{|F(x, \cdot)|_1 |F(\cdot, y)|_1} \right\} dx\, dy.$$

Put

$$A_i = \big\{ (y, \omega) \colon i\lambda(y) \le |F(X_1(\omega), y)| < (i + 1)\lambda(y) \big\}.$$

After interchanging orders of summation, we may express the left side of (3.11) as I + II, where

$$(\mathrm{I}) = \sum_{i=1}^{\infty} \int_0^1 E\left\{ |F(X_1, y)| \log_+\left(\frac{|F(X_1, y)|}{2|F(X_1, \cdot)|_1} \right) \left(\sum_{j=1}^{i \wedge \alpha_0} \frac{1}{j} \right); A_i \right\} dy$$

and

$$(\mathrm{II}) = \sum_{i=1}^{\infty} \int_0^1 E\left\{ |F(X_1, y)| \left(\sum_{j=\alpha_0+1}^{\gamma \wedge i} \frac{1}{j} \log \frac{|F(X_1, y)|}{2j\beta} \right); A_i \right\} dy,$$

with

$$\alpha_0 = \alpha_0(\omega) = \big(|F(X_1, \cdot)|_1 / \beta \big) \vee 1$$

and

$$\gamma = \gamma(\omega, y) = |F(X_1, y)| / 2\beta.$$

It is understood that upper (respectively, lower) limits of summation are to be rounded down (up) to the nearest integer. Now I ~ 0, which means that these terms may be dropped from further consideration. Indeed,

$$E\left\{ \int_0^1 |F(X_1, y)| \log_+\left(\frac{|F(X_1, y)|}{2|F(X_1, \cdot)|_1} \right) dy \log_+\left(\frac{|F(X_1, \cdot)|_1}{\beta} \right) \right\}$$

$$\le E|F(X_1, \cdot)|_{L \log_+ L} \log_+ \frac{|F(X_1, \cdot)|_{L \log_+ L}}{\beta}$$

$$\le \beta.$$

For the estimation of (II), first note that we may replace

$$\sum_{j=\alpha_0}^{\gamma\wedge i}\frac{1}{j}\log\frac{|F(X_1,y)|}{2\beta j}$$

by

$$\int_{\alpha_0}^{\gamma\wedge i\vee\alpha_0}\frac{1}{x}\log\frac{|F(X_1,y)|}{2\beta x}\,dx=\frac{1}{2}\log_+^2\left(\frac{|F(X_1,y)|}{2\beta\alpha_0}\right)-\frac{1}{2}\log_+^2\left(\frac{|F(X_1,y)|}{2\beta(\gamma\wedge i\vee\alpha_0)}\right).$$

This may be seen by recalling that $\alpha_0\geq 1$ and using the Riemann sum upper and lower bounds to the integral. This involves adding or deleting a term from the preceding sum, but such single terms, when substituted into II, produce expressions dominated by β.

Note that the condition $\gamma\wedge i\geq\alpha_0$ entails the condition

$$\frac{|F(X_1,y)|}{\lambda(y)}\geq\frac{|F(X_1,\cdot)|_1}{\beta}$$

on A_i. Also, we have that

$$\int_0^1 E|F(X_1,y)|\log_+^2\frac{\lambda(y)}{\beta}\,dy\sim 0,\tag{3.12}$$

since, by (3.4) applied to one factor of $\log_+$,

$$\begin{aligned}&\int_0^1 E|F(X_1,y)|\log_+^2\frac{\lambda(y)}{\beta}\,dy\\&\quad\leq\int_0^1\left(E|F(X_1,y)|\log_+\frac{|F(X_1,y)|}{\beta}\right)\log_+\frac{\lambda(y)}{\beta}\,dy\\&\qquad+\frac{1}{e}\int_0^1\lambda(y)\log_+\frac{\lambda(y)}{\beta}\,dy\\&\quad\leq\left(1+\frac{1}{e}\right)\beta.\end{aligned}$$

But on A_i we have

$$\log_+^2\frac{|F(X_1,y)|}{2\beta(\gamma\wedge i)}\leq\log_+^2\frac{\lambda(y)}{\beta};$$

hence, the negative $\log_+^2$ term may be dropped from further consideration.

We now have

$$\sum_{i=1}^{\infty}\int_0^1 E\left(|F(X_1,y)|\log_+^2\left(\frac{|F(X_1,y)|}{2\beta\alpha_0}\right);A_i\right)dy=(a)+(b),$$

where

$$(a) = \int_0^1 E\left\{|F(X_1, y)|\log_+^2 \frac{|F(X_1, y)|}{2\beta}; \frac{|F(X_1, \cdot)|_1}{\beta} \le 1 < \frac{|F(X_1, y)|}{\lambda(y)}\right\} dy$$

and

$$(b) = \int_0^1 E\left\{|F(X_1, y)|\log_+^2 \frac{|F(X_1, y)|}{2|F(X_1, \cdot)|_1}; 1 < \frac{|F(X_1, \cdot)|_1}{\beta} \le \frac{|F(X_1, y)|}{\lambda(y)}\right\} dy.$$

Using (3.3) and (3.12), we have

$$(a) \sim \int_0^1 E\left(|F(X_1, y)|\log_+^2 \frac{|F(X_1, y)|}{\lambda(y)}; |F(X_1, \cdot)| \le \beta\right) dy$$

$$\sim \int_0^1 E\left(|F(X_1, y)|\log_+^2 \frac{\beta|F(X_1, y)|}{\lambda(y)|F(X_1, \cdot)|_1}; |F(X_1, \cdot)| \le \beta\right) dy,$$

since, by (3.3),

$$E|F(X_1, \cdot)|_1 \log_+^2 \frac{\beta}{|F(X_1, \cdot)|_1} \le \frac{\beta}{e^2}.$$

Similarly,

$$(b) \sim \int_0^1 E\left\{|F(X_1, y)|\log_+^2 \frac{\beta|F(X_1, y)|}{\lambda(y)|F(X_1, \cdot)|_1}; \right.$$

$$\left. 1 < \frac{|F(X_1, \cdot)|_1}{\beta} \le \frac{|F(X_1, y)|}{\lambda(y)}\right\} dy$$

$$= \int_0^1 E\left(|F(X_1, y)|\log_+^2 \frac{\beta|F(X_1, y)|}{\lambda(y)|F(X_1, \cdot)|_1}; \beta < |F(X_1, \cdot)|_1\right) dy.$$

Thus,

$$(a) + (b) \sim \int_0^1 E|F(X_1, y)|\log_+^2\left(\frac{\beta|F(X_1, y)|}{\lambda(y)|F(X_1, \cdot)|_1}\right) dy.$$

The proof is now completed by using (3.3) and the estimate

$$E\int_0^1 |F(X_1, y)|\log_+^2\left(\frac{\beta|F(\cdot, y)|_1}{\lambda(y)|F|_1}\right) dy \le |F|_1 \log_+^2 \frac{\beta}{|F|_1} \le \frac{\beta}{e^2}. \qquad \square$$

4. Open questions. It seems quite likely that the methods of this paper would extend to the case of triply indexed families of random variables, but the additional complications may be formidable. Also likely is that the family $(F(X_i, Y_j))$ may be replaced by the family $(F(X_i, X_j),\ i \ne j)$, which exhibits a more complicated dependence structure, with the same results as in Theorems 1 and 2 holding. As evidence for this, it is known [7] that the condition

$$\sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\varepsilon_i \varepsilon_j}{ij} F(X_i, X_j) \quad \text{converges a.s. as } n \wedge m \to \infty$$

is equivalent to (1.4) and that the quantity

$$E\left| \sum_{i \neq j} \frac{\varepsilon_i \varepsilon_j}{ij} F(X_i, X_j) \right|$$

is equivalent in size to $\|F\|$ (combine the result of [16] with our Theorem 2).

In another direction, it is known that for nonnegative f the equivalence $Es^*(f) \simeq |f|_{L \log_+ L}$ extends to the case of general ergodic dynamical systems. More precisely, let T be an invertible ergodic measure preserving transformation of a finite measure space $(\Sigma, \mathscr{M}, \mu)$. Then

$$\int \sup_n \frac{1}{n} \left| \sum_{i=1}^{n} f(T^i x) \right| \mu(dx) \approx |f|_{L \log_+ L}$$

(see, e.g., [4], [5] and the references cited therein). Now suppose S is another such transformation. It is natural to conjecture that the analogous condition $\|F\| < \infty$ is necessary and sufficient for the maximal function

$$\sup_{n, m} \frac{1}{nm} \left| \sum_{i=1}^{n} \sum_{j=1}^{m} F(T^i x, S^j y) \right|$$

to be integrable $\mu \otimes \mu$.

REFERENCES

[1] BONAMI, A. (1970). Etudé des coefficients de Fourier des fonctions de $L_p(G)$. *Ann. Inst. Fourier (Grenoble)* **20** 335–402.

[2] BURKHOLDER, D. L. (1962). Successive conditional expectations of an integrable function. *Ann. Math. Statist.* **33** 887–893.

[3] CAMBANIS, S., ROSINSKI, J. and WOYCZINSKI, W. (1985). Convergence of quadratic forms in p-stable random variables and θ_p-radonifying operators. *Ann. Probab.* **13** 885–897.

[4] DANG-NGOC, N. (1980). On the integrability of the maximal ergodic function. *Proc. Amer. Math. Soc.* **79** 565–570.

[5] DAVIS, B. (1982). On the integrability of the ergodic maximal function. *Studia Math.* **73** 153–167.

[6] GUT, A. (1978). Moments of the maximum of normed partial sums of random variables with multidimensional indices. *Z. Wahrsch. verw. Gebiete* **46** 205–220.

[7] KRAKOWIAK, W. and SZULGA, J. (1985). Random multilinear forms. *Ann. Probab.* **14** 955–973.

[8] LEPAGE, R., WOODROOFE, M. and ZINN, J. (1981). Convergence to a stable distribution via order statistics. *Ann. Probab.* **9** 624–632.

[9] MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Sur les fonctions indépendantes. *Fund. Math.* **29** 60–90.

[10] MCCONNELL, T. and TAQQU, M. (1986). Dyadic approximation of double integrals with respect to symmetric stable processes. *Stochastic Process. Appl.* **22** 323–331.

[11] McConnell, T. and Taqqu, M. (1986). Decoupling inequalities for multilinear forms in independent symmetric random variables. *Ann. Probab.* **14** 943–954.

[12] Rosinski, J. (1986). On stochastic integral representation of stable processes with sample paths in Banach spaces. *J. Multivariate Anal.* **20** 277–302.

[13] Rosinski, J. and Woyczinski, W. A. (1984). Moment inequalities for real and vector p-stable stochastic integrals. *Lecture Notes in Math.* **1153** 369–386. Springer, Berlin.

[14] Rosinski, J. and Woyczinski, W. A. (1986). On Itô stochastic integration with respect to p-stable motion: Inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab.* **14** 271–286.

[15] Smythe, R. T. (1974). Sums of independent random variables on partially ordered sets. *Ann. Probab.* **2** 906–917.

[16] Zinn, J. (1985). Comparison of martingale difference sequences. *Lecture Notes in Math.* **1153** 453–457 Springer, Berlin.

Department of Mathematics
200 Carnegie
Syracuse University
Syracuse, New York 13244-1150