convex set in $\mathbf{R}^n$ of volume 1 then for every 1-codimensional subspace H of $\mathbf{R}^n$, $|H \cap C| > \delta$.

3) There is a constant M such that if C is an isotropic symmetric convex set in $\mathbf{R}^n$ of volume 1 then

$$m(C \cap B(M\sqrt{n})) \geq \tfrac{1}{2},$$

where $B(M\sqrt{n})$ is the Euclidean ball of radius $M\sqrt{n}$.

4) There is a constant M such that for every symmetric convex set C in $\mathbf{R}^n$ of volume 1 there is an ellipsoid $\mathcal{E}$ of volume at most M^n such that

$$|\mathcal{E} \cap C| \geq \tfrac{1}{2}.$$

We may remark that such bounds do hold uniformly for the unit balls of spaces with a 1-unconditional basis. This follows from the observation that such a space can be represented on $\mathbf{R}^n$ with an isotropic unit ball C, say, and with the unconditional basis vectors orthogonal. In this situation the section of C perpendicular to a basis vector is also the projection of C onto the orthogonal complement of that vector.

Acknowledgement. This work will form part of a doctoral thesis being written under the supervision of Dr. B. Bollobás whose advice and encouragement have been invaluable.

References

[1] K. M. Ball, *Cube slicing in $\mathbf{R}^n$*, Proc. Amer. Math. Soc. 97 (3) (1986), 465–473.

[2] J. Bourgain, *On high dimensional maximal functions associated to convex bodies*, Amer. J. Math. 108 (1986), 1467–1476.

[3] D. Hensley, *Slicing the cube in $\mathbf{R}^n$ and probability*, Proc. Amer. Math. Soc. 73 (1) (1979), 95–100.

[4] —, *Slicing convex bodies — bounds for slice area in terms of the body's covariance*, ibid. 79 (4) (1980), 619–625.

[5] L. Leindler, *On a certain converse of Hölder's inequality. II*, Acta Sci. Math. (Szeged) 33 (1972), 217–223.

[6] A. Prékopa, *Logarithmic concave measures with application to stochastic programming*, ibid. 32 (1971), 301–316.

[7] J. D. Vaaler, *A geometric inequality with applications to linear forms*, Pacific J. Math. 83 (2) (1979), 543–553.

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS
CAMBRIDGE UNIVERSITY
Mill Lane, Cambridge CB2 1SB, England

Received May 28, 1986
Revised version November 4, 1986 (2172)

STUDIA MATHEMATICA, T. LXXXVIII. (1988)

On the strong maximal function and rearrangements

by

TERRY R. McCONNELL* (Syracuse, N.Y.)

Abstract. We provide sufficient conditions for almost everywhere finiteness, integrability and membership in weak L^1 of the strong maximal function on $\mathbf{T}^2$. These are the weakest possible conditions which are invariant under all measure-preserving transformations of $\mathbf{T}^2$ which preserve the product structure. We also give examples showing that the conditions are not necessary.

1. Introduction. There are many points of contact between probability theory and harmonic analysis. One of the more striking concerns the connections between the Hardy–Littlewood maximal operator and its probabilistic counterpart. In this paper we explore similar connections between the strong maximal operator and a two-parameter probabilistic maximal operator. The differences between the two maximal operators are related to their behavior relative to rearrangements which preserve the product structure.

Let $X_1, X_2, \ldots$ be independent and identically distributed (i.i.d.) random variables on some probability space $(\Omega, \mathfrak{F}, P)$. Suppose also each X_i has a uniform on $[0, 1)$ $(U(0, 1))$ distribution. For Borel functions f on $[0, 1)$ let

$$s_n(f) = \sum_{i=1}^{n} f(X_i), \qquad s^*(f) = \sup_{1 \leq n < \infty} (|s_n(f)|/n).$$

Then by classical results of Khinchin and Kolmogorov and of Marcinkiewicz and Zygmund we have

$$s^*(f) < \infty \text{ a.s.} \quad \text{if and only if} \quad \|f\|_1 < \infty, \tag{1.1}$$

and

$$Es^*(f) \approx |f|_{L\log_+ L} := \int_0^1 |f(x)| \left(1 + \log_+ \frac{|f(x)|}{\|f(x)\|_1}\right) dx. \tag{1.2}$$

---

* Supported in part by NSF grant DMS 85-03775 and by the Alfred P. Sloan foundation.
AMS (1980) *Subject Classification*: Primary 42B25, Secondary 28A15.
Key words and phrases: strong maximal function, differentiation of integrals, Fourier analysis on product domains.

(Here, and throughout the rest of this paper, two extended real quantities A and B are related by $\approx$ and said to be *comparable* provided A is finite if and only if B is, in which case their ratio is bounded below and above by positive absolute constants. The expression $|\ |_{L\log_+ L}$ is not a norm, but it is easily shown to be comparable to any Orlicz-type norm on the Zygmund class $L\log_+ L$.)

These are analogous to the results of Hardy and Littlewood [3] and Stein [11] for the Hardy–Littlewood maximal function $\underline{m}(f)$:

$$(1.3)\qquad \underline{m}(f) < \infty \text{ a.e.} \quad \text{if and only if} \quad \|f\|_1 < \infty,$$

and

$$(1.4)\qquad \|\underline{m}(f)\|_1 \approx |f|_{L\log_+ L}.$$

Here, to be precise, we assume that f is a function on the circle S^1, which we identify as usual with $[0, 1)$. One may complete the analogy suggested by these inequalities by viewing the strong law of large numbers as a counterpart of the Lebesgue differentiation theorem.

Recently there has been some progress in understanding a probabilistic analogue of the two-parameter strong maximal function. Let $\{Y_j\}$ be an independent copy of $\{X_i\}$ and for f Borel on $[0, 1)^2$ let

$$S_{n,m}(f) = \sum_{i=1}^{n}\sum_{j=1}^{m} f(X_i, Y_j), \qquad {}^*S^*(f) = \sup_{1\le n,m<\infty} (|S_{n,m}(f)|/(nm)).$$

The classical conditions $\|f\|_1 < \infty$ and $|f|_{L\log_+ L} < \infty$ are no longer sufficient for finiteness and integrability of ${}^*S^*(f)$. Let

$$\delta_0(f) = \int_{[0,1)^2} |f(x, y)|\left(1+\log_+ \frac{\|f\|_1\,|f(x, y)|}{\|f(x, \cdot)\|_1\,\|f(\cdot, y)\|_1}\right)dx\,dy,$$

$$\Delta(f) = \mathbf{|}|f|\mathbf{|}_{L\log_+ L} + \int_{[0,1)^2} |f(x, y)|\log_+^2 \frac{\|f\|_1\,|f(x, y)|}{\|f(x, \cdot)\|_1\,\|f(\cdot, y)\|_1}\,dx\,dy,$$

where $\mathbf{|}|f|\mathbf{|}_{L\log_+ L} = |F|_{L\log_+ L} + |G|_{L\log_+ L}$ and $F(x) = |f(x, \cdot)|_{L\log_+ L}$, $G(y) = |f(\cdot, y)|_{L\log_+ L}$. Then [6]

$$(1.5)\qquad {}^*S^*(f) < \infty \text{ a.s.} \quad \text{if and only if} \quad \delta_0(f) < \infty,$$

and

$$(1.6)\qquad E^*S^*(f) \approx \Delta(f).$$

The quantities δ_0 and Δ are scalar homogeneous but they are not norms. They are, however, quasi-norms, i.e. satisfy

$$(1.7)\qquad \begin{aligned} &\delta_0(f+g) \leqslant \alpha\delta_0(f) + \alpha\delta_0(g),\\ &\Delta(f+g) \leqslant \alpha\Delta(f) + \alpha\Delta(g), \quad \text{for some } \alpha \geqslant 1. \end{aligned}$$

(See [8, Section 2] for a proof.) The two terms in the definition of Δ are not comparable.

One remarkable property of s^* not shared by ${}^*S^*$ is that $s^*(f)$ finite a.e. already implies that $s^*(f)$ belongs to weak L^1. Weak L^1 (WL^1) is the space of random variables Z having $\|Z\|_{1,\infty} < \infty$ where the quasi-norm $\|\ \|_{1,\infty}$ is defined by

$$\|Z\|_{1,\infty} = \sup_{0<\lambda<\infty} \lambda P(|Z| > \lambda).$$

Indeed, one has

$$(1.8)\qquad \|s^*(f)\|_{1,\infty} \approx \|f\|_1 \approx \|\underline{m}(f)\|_{1,\infty}.$$

In contrast, we may have ${}^*S^*(f) < \infty$ a.s., and yet $\|{}^*S^*(f)\|_{1,\infty} = +\infty$. In [8] we found a necessary and sufficient condition to have ${}^*S^*(f)$ in WL^1. Let $\mu(y) = \|f(\cdot, y)\|_1$, $\lambda(x) = \|f(x, \cdot)\|_1$, $\Lambda = \|f\|_1$, and

$$\delta_*(f) = \sup_{\lambda} \int_0^1\int_0^1 |f(x, y)|\left(1+\log_+ \frac{|f(x, y)|\,\lambda}{(\lambda(x)\vee\lambda)(\mu(y)\vee\lambda)}\right)dx\,dy.$$

Then [8]

$$(1.9)\qquad \|{}^*S^*(f)\|_{1,\infty} \approx \delta_*(f).$$

The counterpart of ${}^*S^*(f)$ in harmonic analysis is the strong maximal function $\underline{M}(f)$ defined by

$$(1.10)\qquad \underline{M}(f)(x_0, y_0) = \sup \frac{1}{|I|\,|J|} \int_{I\times J} |f(x, y)|\,dx\,dy,$$

$(x_0, y_0) \in \boldsymbol{T}^2 = [0, 1) \times [0, 1)$, where the supremum extends over all intervals I containing x_0 and J containing y_0. While all the results of this paper are phrased in the context of functions on $\boldsymbol{T}^2$ or doubly periodic functions, they easily yield results concerning almost everywhere finiteness, local integrability, and local membership in WL^1 for the strong maximal function on $\boldsymbol{R}^2$ as it is usually defined. (For integrability properties of this and related maximal functions over general sets of finite measure see [1].) We should also mention that the strong maximal function is sometimes defined with the additional requirement that the intervals I and J in (1.10) be centered at x_0 and y_0. The same results below hold with either definition.

The main result of this paper is that conditions $\delta_0(f) < \infty$, $\Delta(f) < \infty$ and $\delta_*(f) < \infty$ are sufficient, respectively, for almost everywhere finiteness of $\underline{M}(f)$, and for its membership in L^1 and in WL^1. Unfortunately, while each of these conditions is strictly weaker than any previously known sufficient condition, apparently none of them are necessary conditions. (We thank Carl Mueller for pointing this out. An example, based on his sketch, is given in Section 3 below.) On the other hand, they do come very close to being necessary. Indeed, we make precise below the statement that the conditions given are the weakest for their respective conclusions which are invariant under measure-preserving rearrangements of T^2 which preserve the factors. (The classical one-parameter conditions are invariant under all measure-preserving rearrangements.)

We now turn to the precise statements.

Let $\mathfrak{F}_n$ be the nth dyadic σ-field of subsets of $[0, 1)$, i.e., $\mathfrak{F}_n$ is generated by $I_{k,n} = [k2^{-n}, (k+1)2^{-n})$, $k = 0, 1, \ldots, 2^n - 1$. We denote by $[\mathfrak{F}_n]$ the collection of all the $I_{k,n}$.

Let $\mathfrak{D}$ denote the class of all one-to-one measure-preserving maps of $[0, 1)$ onto itself which, for all sufficiently large n, induce a permutation of $[\mathfrak{F}_n]$. For a pair $T, S \in \mathfrak{D}$ we denote by $f \circ T, S$ the composite function $f(T(x), S(y))$.

THEOREM 1.1. *Let f belong to $L^1(T^2)$. Then $\delta_0(f) < \infty$ implies $\underline{M}(f) < \infty$ a.e. Indeed, given $\varepsilon > 0$ there is $\delta > 0$ such that*

$$\delta_0(f) < \delta \Rightarrow |\{\underline{M}(f) > \varepsilon\}| < \varepsilon. \tag{1.11}$$

Also, we have

$$\|\underline{M}(f)\|_{1,\infty} \leqslant c\delta_*(f), \tag{1.12}$$

$$\|\underline{M}(f)\|_1 \leqslant c\Delta(f) \tag{1.13}$$

for some constant c.

In [5] (see also [9]) it is shown that dyadic functions are dense in the space determined by $\delta_0(f) < \infty$. We give a simple proof below in Section 3. Combining this with (1.11), standard arguments yield

COROLLARY 1.1. *The condition $\delta_0(f) < \infty$ is sufficient for strong differentiability of $\int_{T^2} f(x, y)\,dx\,dy$.*

(See [2] for background on differentiation of integrals.)

THEOREM 1.2. *For each $\lambda > 0$ we have*

$$P(*S^*(f) > \lambda) \leqslant \sup_{T,S \in \mathfrak{D}} |\{\underline{M}(f \circ T, S) > \lambda\}|. \tag{1.14}$$

Also,

$$c\Delta(f) \leqslant \sup_{T,S} \|\underline{M}(f \circ T, S)\|_1, \tag{1.15}$$

$$c\delta_*(f) \leqslant \sup_{T,S} \|\underline{M}(f \circ T, S)\|_{1,\infty}. \tag{1.16}$$

Throughout the paper the letter c denotes a constant which may change from line to line.

We shall prove Theorems 1.1 and 1.2 in Section 2. In Section 3 we present examples which show that the suprema in (1.14) and (1.15) cannot be omitted.

We currently have the means to extend results (1.11) and (1.14) to 3 parameters. The analogue of the condition $\delta_0(f) < \infty$ is the condition given in [7] for triple integrability of f with respect to the symmetric Cauchy process. (Take $\alpha = 1$ in Theorem 1.1 of [7]. See Section 2 of [6] and the references quoted there for connections between maximal functions and Cauchy processes.) Unfortunately, this condition is vastly more complicated than $\delta_0(f) < \infty$ and so it appears to be very difficult, if not impossible, to extend the results of the present paper to n parameters.

2. Proofs. To prove Theorem 1.1 we shall need some further one-parameter results. These are summarized in the following lemma (see [8, Section 2] for proofs or references).

LEMMA 2.1. *For each Borel function f on $[0, 1)$ we have $\sup_i(|f(X_i)|/i) < \infty$ if and only if $\|f\|_1 < \infty$. Moreover,*

$$\|\sup_i(|f(X_i)|/i)\|_1 \approx |f|_{L\log_+ L}. \tag{2.1}$$

For $\lambda > 6\|f\|_1$ we have

$$\begin{aligned} \tfrac{1}{2}E(|f(X_1)|;\ |f(X_1)| > \lambda) &\leqslant \lambda P(s^*(|f|) > \lambda) \\ &\leqslant 2e\lambda P\big(\sup_i(|f(X_i)|/i) > \lambda\big) \\ &\leqslant 4eE(|f(X_1)|;\ |f(X_1)| > \lambda/2). \end{aligned} \tag{2.2}$$

Finally,

$$P\big(\sup_i(|f(X_i)|/i) > \tfrac{1}{4}\|f\|_1\big) \geqslant \tfrac{1}{2}. \tag{2.3}$$

Now let f be a function such that $\delta_0(f) < \infty$. Then $*S^*(|f|) < \infty$ a.s., and hence, by Fubini's theorem and the strong law of large numbers,

$$\int_0^1 \sup_i(|f(X_i, y)|/i)\,dy < \infty \quad \text{a.s.}$$

Thus, from (1.3) and an obvious pointwise inequality we conclude

$$\sup_i [m_2(f(X_i, \cdot))/i] < \infty \quad \text{a.s., a.e.,}$$

where $m_2(f(x, \cdot))$ denotes the Hardy–Littlewood maximal function (one-dimensional) of f considered as a function of its second variable. Repeating this argument, we conclude, in turn,

$$\int_0^1 m_2(f(x, \cdot))\,dx \leqslant \infty \quad \text{a.e.,}$$

$m_1(m_2(f)) < \infty$ a.e., and hence $\underline{M}(f) < \infty$ a.e. The inequality (1.11) is now easily proved by an indirect argument as in [8].

The proof of (1.13) is similar, using (1.2) in place of (1.1), etc. One merely trades, twice, a probabilistic for a Hardy–Littlewood maximal function.

The weak-type inequality (1.12) is somewhat more difficult to obtain. We combine (2.2) and (2.3) with the Hardy–Littlewood inequality,

$$\lambda\,|\{\underline{m}(f) > \lambda\}| \leqslant 2 \int_{\{|f| > \lambda/2\}} |f|. \tag{2.4}$$

We have

$$\|m_1(m_2(f))\|_{1,\infty} \leqslant 12 \Big\|\int_0^1 m_2(f(x,\cdot))\,dx\Big\|_{1,\infty} + \sup_{0<\lambda<\infty} \lambda \int_0^1 |\{x\colon m_1(m_2 f)(x, y) > \lambda\}|\, I\{\lambda > 12 \int_0^1 m_2(f(x', y))\,dx'\}\,dy. \tag{2.5}$$

For the first term we estimate, using (2.3),

$$\begin{aligned}\lambda\,\Big|\Big\{y\colon \int_0^1 m_2(f(x, y))\,dx > \lambda\Big\}\Big| &\leqslant 2\lambda E\,\Big|\Big\{y\colon \sup_i m_2\Big(\frac{f(X_i, y)}{i}\Big) > \frac{\lambda}{4}\Big\}\Big| \\ &\leqslant 2\lambda E\,\Big|\Big\{y\colon m_2\Big(\sup_i \frac{|f(X_i, y)|}{i}\Big) > \frac{\lambda}{4}\Big\}\Big|.\end{aligned}$$

Thus we obtain

$$\Big\|\int_0^1 m_2(f(x,\cdot))\,dx\Big\|_{1,\infty} \leqslant 48 \Big\|\int_0^1 \sup_i \frac{|f(X_i, y)|}{i}\,dy\Big\|_{1,\infty} + 8 \sup_{0<\lambda<\infty} \lambda E\,\Big|\Big\{y\colon m_2\Big(\sup_i \frac{|f(X_i, y)|}{i}\Big) > \lambda,\ \lambda > 6\int_0^1 \sup_i \frac{|f(X_i, y')|}{i}\,dy'\Big\}\Big|. \tag{2.6}$$

For the first term here, we have by (2.3),

$$\Big\|\int_0^1 \sup_i \frac{|f(X_i, y)|}{i}\,dy\Big\|_{1,\infty} \leqslant 8 \Big\|\sup_{j,i} \frac{|f(X_i, Y_j)|}{ij}\Big\|_{1,\infty}.$$

Thus we obtain

$$\Big\|\int_0^1 \sup_i \frac{|f(X_i, y)|}{i}\,dy\Big\|_{1,\infty} \leqslant c\delta_*(f) \tag{2.7}$$

by applying the equivalence

$$\Big\|\sup_{i,j} \frac{|f(X_i, Y_j)|}{ij}\Big\|_{1,\infty} \approx \delta_*(f) \tag{2.8}$$

of [8]. What remains is to obtain an inequality similar to (2.7) for the second terms of (2.5) and (2.6).

Combining (2.4) with (2.2) we see that the second term in (2.5) is bounded by a multiple of

$$\begin{aligned}\Big\|\sup_i \frac{m_2(f(X_i, y))}{i}\Big\|_{1,\infty} &\leqslant \Big\|m_2\Big(\sup_i \frac{|f(X_i, y)|}{i}\Big)\Big\|_{1,\infty} \\ &\leqslant 12 \Big\|\int_0^1 \sup_i \frac{|f(X_i, y)|}{i}\,dy\Big\|_{1,\infty} + \sup_{0<\lambda<\infty} \lambda E\,\Big|\Big\{y\colon m_2\Big(\sup_i \frac{|f(X_i, y)|}{i}\Big) > \lambda, \\ &\qquad \lambda > 12 \int_0^1 \sup_i \frac{|f(X_i, y')|}{i}\,dy'\Big\}\Big|\end{aligned}$$

which is essentially the same as (2.6), so all that remains is to estimate the second term here. For this, we again combine (2.4) and (2.2) and apply (2.8) as previously. The proof of Theorem 1.1 is complete.

Remark 2.1. An examination of the proof shows that we have actually proved a stronger theorem with the strong maximal function $\underline{M}$ being replaced by the *iterated maximal function* $m_2(m_1(f))$. Thus the methods of this paper cannot distinguish amongst $m_1 m_2$, $m_2 m_1$ and $\underline{M}$. Since these operators do behave differently (see Remark 3.1), it is an open problem to devise function space quasi-norms which can make these distinctions.

We now turn to the proof of Theorem 1.2. It will be convenient to replace $\underline{M}$ here by a somewhat more tractable object. Recall that the nth dyadic σ-field $\mathfrak{F}_n$ is generated by the collection $[\mathfrak{F}_n]$ of 2^n intervals $I_{k,n}$ each having length 2^{-n}. Define marginal conditional expectations by

$$E(f \mid \mathfrak{F}_n^1)(x, y) = \sum_{i=0}^{2^n-1} \Big(2^n \int_{I_{i,n}} f(x', y)\,dx'\Big) 1_{I_{i,n}}(x)$$

with an analogous expression for $E(f \mid \mathfrak{F}_n^2)$. Then

$$E(E(f \mid \mathfrak{F}_n^1) \mid \mathfrak{F}_m^2) = E(f \mid \mathfrak{F}_n^1 \otimes \mathfrak{F}_m^2).$$

The family

$$f_{n,m} = E(|f| \mid \mathfrak{F}_n^1 \otimes \mathfrak{F}_m^2) \tag{2.9}$$

is a two-parameter martingale. Its maximal function, the *dyadic maximal function* $M(f)$ (no underscore), is closely related to $\underline{M}(f)$. We will need the following technical result about M. (Note that $\underline{M}$ dominates M pointwise.)

Lemma 2.2. *Suppose $\|f\|_\infty < \infty$. Then there exists a sequence f_n of dyadic functions such that f_n converges to f a.e., $M(f_n)$ converges to $M(f)$ a.e.,*

$$\sup_{S,T\in\mathfrak{T}} \|M(f_n \circ T, S) - M(f \circ T, S)\|_{1,\infty} \to 0 \tag{2.10}$$

as $n \to \infty$, and

$$\sup_{T,S} \int_{T^2} |M(f_n \circ T, S) - M(f \circ T, S)| \to 0 \tag{2.11}$$

as $n \to \infty$.

Proof. In the notation (2.9) above we take $f_n = f_{n,n}$ and note that, by the martingale convergence theorem, $f_n \to f$ a.e. and in each L^p, $0 \leqslant p < \infty$. Thus $f_n \circ T, S \to f \circ T, S$ almost everywhere for each pair T, S in $\mathfrak{D}$ and in all such L^p.

Now for each Borel g we have [4]

$$\left(\int M(g)^2\right)^{1/2} \leqslant C\|g\|_2.$$

But, for any $\lambda > \varepsilon > 0$, and T, S in $\mathfrak{D}$,

$$\lambda\left|\{|M(f_n \circ T, S) - M(f \circ T, S)| > \lambda\}\right| \leqslant \varepsilon^{-1}\|M(f_n \circ T, S) - M(f \circ T, S)\|_2^2$$
$$\leqslant \varepsilon^{-1}\|M((f_n - f) \circ T, S)\|_2^2 \leqslant c\varepsilon^{-1}\|(f_n - f) \circ T, S\|_2^2 = c\varepsilon^{-1}\|f_n - f\|_2^2.$$

Taking the supremum over such $\lambda \geqslant \varepsilon$, T and S we obtain

$$\overline{\lim_{n\to\infty}} \sup_{T,S} \|M(f_n \circ T, S) - M(f \circ T, S)\|_{1,\infty} \leqslant \varepsilon.$$

This implies (2.10). The proof of (2.11) is similar.

The next proposition is an analogue of Fatou's lemma for the quasi-norms δ_0, δ_* and Δ. The sequence f_n is that of Lemma 2.2.

Proposition 2.1. *We have*

$$\text{(2.12)} \quad \delta_*(f) \leqslant \underline{\lim_{n\to\infty}}\, \delta_*(f_n), \quad \Delta(f) \leqslant \underline{\lim_{n\to\infty}}\, \Delta(f_n), \quad \delta_0(f) \leqslant \underline{\lim_{n\to\infty}}\, \delta_0(f_n),$$

$$\text{(2.13)} \quad \delta_*(f) \leqslant \underline{\lim_{N\to\infty}}\, \delta_i(|f| \wedge N), \quad \Delta(f) \leqslant \underline{\lim_{N\to\infty}}\, \Delta(|f| \wedge N),$$
$$\delta_0(f) \leqslant \underline{\lim_{N\to\infty}}\, \delta_0(|f| \wedge N).$$

Proof. We have $\|f_n\|_1 \to \|f\|_1$ and $|f_n|_{L\log_+L} \to |f|_{L\log_+L}$ as extended real numbers, thus, by Fubini's theorem and Fatou's lemma,

$$\|f_n(x, \cdot)\|_1 \to \|f(x, \cdot)\|_1 \quad \text{and} \quad |f_n(x, \cdot)|_{L\log_+L} \to |f(x, \cdot)|_{L\log_+L}$$

for almost every x, with a similar result for the other variable. All the statements of the lemma are now easy consequences of these observations and Fatou's lemma.

Recall that a one-to-one, onto, measure-preserving map σ: $[0, 1) \to [0, 1)$ is called a *dyadic map* if, for all sufficiently large n, σ induces a permutation of $[\mathfrak{F}_n]$. The collection of all such maps is denoted by $\mathfrak{D}$.

The next result is the main step in the proof of Theorem 1.2.

Lemma 2.3. *For each fixed pair of integers N and M there is a dyadic map σ such that for each f measurable $\mathfrak{F}_M \otimes \mathfrak{F}_M$*

$$\text{(2.14)} \quad P\left(\max_{1\leqslant n,m\leqslant N} \frac{1}{nm}\left|\sum_{i,j=1}^{n,m} f(X_i, Y_j)\right| > \lambda\right) \leqslant |\{M(f \circ \sigma^{-1}, \sigma^{-1}) > \lambda\}|, \quad \lambda > 0.$$

We shall prove Lemma 2.3 below after having first shown how Theorem 1.2 follows.

First note that we may assume f nonnegative, and that it is enough to prove (1.14)–(1.16) with M replacing $\underline{M}$ since the latter is pointwise larger. Inequality (1.14) for dyadic f now follows from (2.14) after passage to the limit as $N \to \infty$. Suppose that f is bounded and let f_n be as in Lemma 2.2. It follows easily from Fatou's lemma that

$$\text{(2.15)} \quad P({}^*S^*(f) > \lambda) \leqslant \underline{\lim_{n\to\infty}}\, P({}^*S^*(f_n) > \lambda)$$

(${}^*S^*(f)$ is the a.s. limit of the ${}^*S^*(f_n)$—use the L^2 maximal inequality, for example.) To complete the proof of (1.14) for bounded f it is enough to show that

$$\text{(2.16)} \quad \underline{\lim_{n\to\infty}} \sup_{T,S} |\{M(f_n \circ T, S) > \lambda\}| \leqslant \sup_{T,S} |\{M(f \circ T, S) > \lambda\}|.$$

Fix a $\lambda > 0$. We may assume (Prop. 3.1) $\delta_0(|f_n - f|) \to 0$ after, perhaps, passing to a subsequence. Given $\varepsilon > 0$ choose δ as in (1.11) so that $\delta_0(g) < \delta$ implies

$$\sup_{T,S\in\mathfrak{D}} |\{M(g \circ T, S) > \varepsilon\}| < \varepsilon.$$

(The estimate of (1.11) is independent of T, S because the family $\{T(X_i), S(Y_j)\}$ has the same distribution as the family $\{X_i, Y_j\}$. See the proof of Theorem 1.1.) The desired result now follows easily from

$$\sup_{T,S} |\{M(f_n \circ T, S) > \lambda\}| \leqslant \sup_{T,S} |\{M(f \circ T, S) > \lambda - \varepsilon\}|$$
$$+ \sup_{T,S} |\{M((f_n - f) \circ T, S) > \varepsilon\}|.$$

The passage from bounded f to general f is easily effected since, for f nonnegative, $M(f \wedge N) \leqslant M(f)$.

The proofs of (1.15) and (1.16) follow the same pattern and use the remaining conclusions of Lemma 2.2 and Proposition 2.1. (They do not, unfortunately, follow directly from (1.14).)

Remark 2.2. An examination of the argument just given shows that one may replace the supremum over $T, S \in \mathfrak{D}$ with a supremum over $T, T \in \mathfrak{D}$. This is apparently related to the recently discovered "decoupling phenomena" (see [10]).

We now turn to the proof of Lemma 2.3. It will be convenient to work with an explicit realization of the variables X_i and Y_j on $[0, 1)$. Let $[\ ]$ denote the greatest integer function and define θ: $[0, 1) \to [0, 1)$ by $\theta(x) = 2x - [2x]$. Fix M in the statement of Lemma 2.3. It is then easy to construct inductively a sequence of integers n_i such that $n_0 = 0$, $n_i < n_{i+1}$ and such that for each g measurable $\mathfrak{F}_M$ the sequence $g(\theta^{n_i}(x))$ is i.i.d. Let $U_i(x)$

$=\theta^{n_i}(x)$. (Note that $U_0(x)=x$.) The U_i each have a $U(0, 1)$ distribution. They are not independent, but we may treat them as if they were, for given any $f\in \mathfrak{F}_M\otimes\mathfrak{F}_M$, the family $\{f(U_i(x), U_j(y))\}$ has the same distribution as the family $\{f(X_i, Y_j)\}$.

For our purposes below it is necessary to consider Banach-valued functions.

LEMMA 2.4. *For each pair of integers M and N there is a dyadic map σ such that*

(2.17)
$$\sup_{1\leqslant n\leqslant N}\Big|n^{-1}\sum_{i=1}^{n} f(U_i(x))\Big|_E \leqslant \sup_{1\leqslant n\leqslant \infty}|E(f\circ\sigma^{-1}\mid \mathfrak{F}_n)|_E(\sigma(x)),\qquad x\in[0, 1),$$

for every Banach space E and E-valued functions f measurable $\mathfrak{F}_M$. (It is vital that σ be independent of E and f.)

Proof. We shall use the L^1-valued function

$$G(x)=\sum_{i=0}^{2^M-1} x_i 1_{I_{i,M}}(x),$$

where $x_i=2^M\chi_{I_{i,M}}$. (χ and 1 denote characteristic functions on distinct copies of $[0, 1)$.) Set

$$s_{-k}(x)=\frac{1}{k+1}\sum_{j=0}^{k} G(U_j(x))\quad \text{for } 0\leqslant k\leqslant N-1,$$

$$s_{-N}(x)=\bar{G}:=\int_0^1 G(x')\,dx'\quad \text{(constant)}.$$

Then the sequence $(s_{-N},\ldots, s_{-1}, s_0)$ is a martingale. To complete the proof, we need:

LEMMA 2.5. *For each pair of integers M and N there are an integer $k(M, N)$, a dyadic function $H\in\mathfrak{F}_{k(M,N)}$, and positive integers $k_1\leqslant\ldots\leqslant k_N=k(M, N)$ such that*

(2.18)
$$(s_{-i},\ i=N,\ldots, 0)\stackrel{L}{=}(\bar{G},\ E(H\mid\mathfrak{F}_{k_j}),\ j=1,\ldots, N).$$

Proof. It is easy to check that there is a large enough L such that each s_{-i} is measurable $\mathfrak{F}_L$, by considering the slopes of the U_i.

We prove more generally that for any martingale $f_0, f_1, \ldots, f_N$ with f_0 constant and each f_k measurable with respect to $\mathfrak{F}_L$ for some L, there is a dyadic function F measurable with respect to $\mathfrak{F}_{LN}$ such that

(2.19)
$$f_0,\ (f_i)_{i=1}^N\stackrel{L}{=}\bar{F},\ (E(F\mid\mathfrak{F}_{Li}))_{i=1}^N.$$

This we do by induction on N.

For $N=1$ we may take $F=f_1$. Suppose the result has been proved for $N-1\geqslant 0$. Note the following property of θ: For any E and E-valued function h on $[0, 1)$, and any set $A\in\mathfrak{F}_L$, the conditional distribution of $1_A\cdot(h\circ\theta^L)$ given A is the same as the (unconditional) distribution of h.

Now let $F_0\in\mathfrak{F}_{LN-L}$ be as constructed for $N-1$. Let $x_1,\ldots, x_k$ be the distinct values assumed by $(\bar{F}_0, E(F_0\mid\mathfrak{F}_L),\ldots, F_0)$ and $A_1,\ldots, A_k$ the respective sets (in $\mathfrak{F}_{LN-L}$) where these values are assumed. Then we have

$$|A_i| = P((f_0,\ldots, f_{N-1})=x_i).$$

Let h_i be $\mathfrak{F}_L$-measurable functions on $[0, 1)$ such that h_i has distribution equal to that of f_N-f_{N-1} given the event $\{(f_0,\ldots, f_{N-1})=x_i\}$. Let

$$F=F_0+\sum_{i=1}^{k} 1_{A_i} h_i\circ\theta^{LN-L}.$$

Then it is easy to check that F has the required properties.

We now return to the proof of Lemma 2.4. An immediate consequence of (2.18) is that the atoms of the σ-field generated by the martingale on the left have the same measure as the corresponding atoms generated by the martingale on the right. Noting that both these martingales are measurable $\mathfrak{F}_{k(M,N)}$ we may then easily construct a dyadic map σ inducing a permutation of $[\mathfrak{F}_{k(M,N)}]$ such that

(2.20)
$$(s_{-N},\ldots, s_0)(x)=(\bar{G},\ E(H\mid\mathfrak{F}_{k_1}),\ldots, E(H\mid\mathfrak{F}_{k_N}))(\sigma(x)).$$

In particular, equality of the last components gives

$$E(H\mid\mathfrak{F}_{k_N})(\sigma(x))=G(U_0(x))=G(x),$$

so by the iteration property of conditional expectation and $\mathfrak{F}_{k_N}=\mathfrak{F}_{k(M,N)}$ we obtain

(2.21)
$$(s_{-N},\ldots, s_0)(x)=(\bar{G},\ E(G\circ\sigma^{-1}\mid\mathfrak{F}_{k_1}),\ldots, G\circ\sigma^{-1}(\sigma(x))).$$

Now let $f\in\mathfrak{F}_M(E)$ and define $T_f\colon L^1(S^1)\to E$ by

$$T_f(h)=\int_{S^1} f(x)h(x)\,dx.$$

Thus

$$(T_f G)(x)=\sum_{i=0}^{2^M-1} T_f(2^M\chi_{I_{i,M}})1_{I_{i,M}}(x)=\sum_{i=0}^{2^M-1}(f|_{I_{i,M}})1_{I_{i,M}}(x)=f(x).$$

Applying T_f to both sides of (2.21) and using the operator linearity of

Bochner-conditional expectations we obtain

$$(\bar{f}, (N-i+1)^{-1}\sum_{j=0}^{N-i} f(U_j(x)),\ i=1,\ldots,N) = (\bar{f}, E(f\circ\sigma^{-1}\mid \mathfrak{F}_{k_i})(\sigma(x)),$$
$$i=1,\ldots,N).$$

The desired conclusion (2.17) follows by comparing the maxima of these sequences and using the inequality

$$\max\{|E(f\circ\sigma^{-1}\mid \mathfrak{F}_{k_i})|_E\colon 1\leqslant i\leqslant N)\leqslant \sup_{1\leqslant n<\infty}|E(f\circ\sigma^{-1}\mid \mathfrak{F}_n)|_E.$$

The proof of Lemma 2.4 is complete.

Now we return to the proof of (2.14). For this purpose, fix M and choose $f\in \mathfrak{F}_M^1\otimes\mathfrak{F}_M^2$. Fix y and introduce the finite-dimensional Banach space

$$E_y = \{(a_0,\ldots,a_{2^M-1})\colon a_i\in \boldsymbol{R}\}$$

with norm

$$|(a_1,\ldots,a_{2^M-1})|_y = m\Big(\sum_{i=0}^{2^M-1} a_i\chi_{I_{i,M}}\Big)(y)$$

(dyadic maximal function evaluated at y). Denote by e_i the standard basis of E_y. Define $g\colon S^1\to E_y$ by

$$g(x) = \sum_{l=0}^{2^M-1}|f(x, l/2^M)|\,e_l.$$

Let σ be as in Lemma 2.4 (for M and the given N). The point of these definitions is that we have

$$M(f\circ\sigma^{-1}, I)(\sigma(x), y) = \sup_{0\leqslant j\leqslant M}|E(g\circ\sigma^{-1}\mid\mathfrak{F}_j^1)|_{E_y}(\sigma(x)),$$

where I denotes the identity. Thus, by Lemma 2.4 we have

$$M(f\circ\sigma^{-1}, I)(\sigma(x), y)\geqslant \sup_{1\leqslant i\leqslant N}\Big|i^{-1}\sum_{\alpha=1}^{i} g(U_\alpha(x))\Big|_{E_y}$$
$$= \sup_{1\leqslant i\leqslant N} m\Big(\sum_{l=0}^{2^M-1} i^{-1}\sum_{\alpha=1}^{i}|f(U_\alpha(x), l/2^M)|\,I_{l,M}(\cdot)\Big)(y)$$
$$= \sup_{1\leqslant i\leqslant N} m(V_i(x,\cdot))(y),\quad \text{say (defining } V_i).$$

Now fix x and allow y to vary. The B-space E_y is no longer relevant. From now on everything is real-valued. Apply Lemma 2.4, this time with $E=\boldsymbol{R}$ to

show

$$m\big(V_i(x,\sigma^{-1}(\cdot))\big)(\sigma(y))\geqslant \sup_{1\leqslant j\leqslant N} j^{-1}\sum_{\beta=1}^{j} V_i(x, U_\beta(y)).$$

Combining this with the above, we conclude

$$M(f\circ\sigma^{-1},\sigma^{-1})(\sigma(x),\sigma(y))\geqslant \sup_{1\leqslant i,j\leqslant N}(ij)^{-1}\sum_{\alpha,\beta=1}^{i,j}|f(U_\alpha(x), U_\beta(y))|.$$

The desired inequality (2.14) follows, and hence the proof of Theorem 1.2 is complete.

3. Complements. C. Mueller has shown us an example of a function f having $\delta_0(f)=+\infty$, but with $\underset{\sim}{M}(f)<\infty$ a.e. Expanding, somewhat, on his idea we present such an example here and at the same time construct a function f having $\Delta(f)=\infty$ but $\underset{\sim}{M}(f)$ in L^1.

Example 3.1. Let h be an even nonnegative integrable function of period 1 which is nonincreasing on $(0, 1/2]$. We begin by proving that if $-1/2\leqslant x_0\leqslant 1/2$, $x_0\neq 0$, and if I is a finite interval of any length containing x_0 then

$$\frac{1}{|I|}\int_I h(t)\,dt\leqslant \frac{4}{|x_0|}\int_0^{|x_0|} h(t)\,dt. \tag{3.1}$$

To see this, let $I^* = I\cup(-I)$. Then

$$\frac{1}{|I|}\int_I h(t)\,dt\leqslant\frac{2}{|I^*|}\int_{I^*} h(t)\,dt.$$

We may assume $0<x_0\leqslant 1/2$ and that x_0 belongs to the interior of I. Now let J denote the set $(x_0, 1-x_0)$ together with all translates of this interval by integers. Since

$$\sup\{h(t)\colon t\in J\}\leqslant\inf\{h(t)\colon t\in I^*\setminus J\}$$

we easily deduce

$$\frac{1}{|I^*|}\int_{I^*} h(t)\,dt\leqslant\frac{2}{|I^*\setminus J|}\int_{I^*\setminus J} h(t)\,dt.$$

Now the set $I^*\setminus J$ consists of a symmetric subset of $[-x_0, x_0]$ containing x_0, together with translates of this set. From the form of this set and of h we clearly have

$$\frac{1}{|(I^*\setminus J)\cap[-x_0,x_0]|}\int_{(I^*\setminus J)\cap[-x_0,x_0]} h(t)\,dt\leqslant\frac{1}{2x_0}\int_{-x_0}^{x_0} h(t)\,dt.$$

The desired inequality follows easily from this. We may easily extend the inequality for any noninteger x_0 as

$$(3.2)\qquad \frac{1}{|I|}\int_I h(t)\,dt \leqslant \frac{4}{\{|x_0|\}}\int_0^{\{|x_0|\}} h(t)\,dt,$$

where $\{|x_0|\}$ denotes the fractional part of $|x_0|$, $\{|x_0|\} = |x_0|-[|x_0|]$.

Now define $f(x, y)$ by $f(x, y) = h(x-y)$.

LEMMA 3.1. *For any* (x_0, y_0) *with* $x_0 \neq y_0$ *we have*

$$(3.3)\qquad \underline{M}(f)(x_0, y_0) \leqslant \frac{16}{\{|x_0-y_0|\}}\int_0^{\{|x_0-y_0|\}} h(t)\,dt.$$

Proof. Let R be any rectangle containing the point (x_0, y_0) in its interior (by a limiting argument we need only consider this case). Let L be the line $x-y = x_0-y_0$ and z_+ and z_- the points where this line meets the boundary of R. Also let L_+ and L_- be lines perpendicular to L through z_+ and z_- respectively. The lines L_+ and L_- divide R into 3 pieces. We denote by T the part between the lines, by T_+ the part above L_+, and by T_- the remaining part. These last two may lie either above or below the line L, depending on the location of (x_0, y_0). We will estimate each of the integrals

$$\int_{T_\pm} f\,dx\,dy,\qquad \int_T f\,dx\,dy.$$

To estimate $\int_{T_-} f\,dx\,dy$ we suppose that T_- lies above L (the other case being similar) and choose a polar coordinate system centered at z_-. Then there is a continuous, piecewise linear function φ such that

$$T_- = \{(r, \theta)\colon\ 3\pi/4 \leqslant \theta \leqslant \pi,\ 0 \leqslant r \leqslant \varphi(\theta)\}.$$

Thus

$$\begin{aligned}\int_{T_-} f\,dx\,dy &= \int_{3\pi/4}^{\pi}\int_0^{\varphi(\theta)/\sqrt{2}} h\big(x_0-y_0+r(\cos\theta-\sin\theta)\big)\,r\,dr\,d\theta\\ &= -\int_{3\pi/4}^{\pi}(\cos\theta-\sin\theta)^{-2}\int_0^{\varphi(\theta)(\cos\theta-\sin\theta)/\sqrt{2}} h(x_0-y_0+t)\,t\,dt\,d\theta\\ &\leqslant \int_{3\pi/4}^{\pi}\int_0^{\sqrt{2}\varphi(\theta)} h(x_0-y_0-t)\,t\,dt\,d\theta.\end{aligned}$$

By the second mean value theorem there is a function $\xi(\theta)$ satisfying $0 \leqslant \xi(\theta) \leqslant \sqrt{2}\varphi(\theta)$ such that

$$\int_0^{\sqrt{2}\varphi(\theta)} h(x_0-y_0-t)\,t\,dt = \xi(\theta)\int_0^{\sqrt{2}\varphi(\theta)} h(x_0-y_0-t)\,dt.$$

Thus, using (3.2), we obtain

$$\int_{T_-} f\,dx\,dy \leqslant \Big(\int_{3\pi/4}^{\pi}\sqrt{2}\,\varphi(\theta)\,\xi(\theta)\,d\theta\Big)\frac{4}{\{|x_0-y_0|\}}\int_0^{\{|x_0-y_0|\}} h(t)\,dt.$$

But

$$\int_{3\pi/4}^{\pi}\sqrt{2}\,\varphi(\theta)\,\xi(\theta)\,d\theta \leqslant 4\int_{3\pi/4}^{\pi}\tfrac{1}{2}\varphi^2(\theta)\,d\theta = 4|T_-|.$$

A similar estimate applies to T_+. Hence

$$(3.4)\qquad \int_{T_\pm} f\,dx\,dy \leqslant \frac{16|T_\pm|}{\{|x_0-y_0|\}}\int_0^{\{|x_0-y_0|\}} h(t)\,dt.$$

To estimate $\int_T f\,dx\,dy$ we make the change of variables $u = (x+y)/2$, $v = (x-y)/2$ and find constants $a < (x_0+y_0)/2 < b$ and piecewise linear functions $\varphi(u)$ and $\psi(u)$ such that $\varphi(u) \leqslant (x_0-y_0)/2 \leqslant \psi(u)$ and

$$T = \{((u+v)/2,\ (u-v)/2)\colon\ \varphi(u) \leqslant v \leqslant \psi(u),\ a \leqslant u \leqslant b\}.$$

Then, by (3.2),

$$\begin{aligned}\int_T f\,dx\,dy &= \int_a^b\int_{\varphi(u)}^{\psi(u)} h(2v)\,dv\,du = \int_a^b \frac{\psi(u)-\varphi(u)}{2\psi(u)-2\varphi(u)}\int_{2\varphi(u)}^{2\psi(u)} h(v)\,dv\,du\\ &\leqslant \frac{8|T|}{\{|x_0-y_0|\}}\int_0^{\{|x_0-y_0|\}} h(t)\,dt.\end{aligned}$$

Combining this with (3.4) yields (3.3) since R was arbitrary.

Now choose h to have all the above properties and in addition such that $|h|_{L\log_+ L} = +\infty$. Defining f as above, we have for each x or y in $[0, 1)$, $\|f(x, \cdot)\|_1 = \|f(\cdot, y)\|_1 = \|h\|_1$. Thus the normalizing factors in the denominator of the logarithm in $\delta_0(f)$ are constant. Since $h \notin L\log_+ L$ we conclude $\delta_0(f) = \infty$. But, by (3.3), we have $\underline{M}(f) < \infty$ a.e.

To obtain an example of a function f having $\Delta(f) = \infty$ but $\underline{M}(f) \in L^1$, we may proceed as above by taking $h \notin L\log_+^2 L$ (which gives $\Delta(f) = \infty$) and $h \in L\log_+ L$. The latter condition, together with (3.3) and Lemma 3.2 below, yields $\underline{M}(f) \in L^1$.

LEMMA 3.2. *Let* h *be nondecreasing and nonnegative on* $(0, 1]$. *Then*

$$\tfrac{1}{2}|h|_{L\log_+ L} \leqslant \int_0^1 x^{-1}\int_0^x h(t)\,dt\,dx \leqslant (2+\ln 2)|h|_{L\log_+ L}.$$

This result (with possibly different constants) is known, but we will present an elementary proof here for the reader's convenience.

Proof. We may assume $\|h\|_1 < \infty$. For each $\lambda > 0$ we have $|\{t: h(t) \geqslant \lambda\}| \leqslant \|h\|_1/\lambda$. Putting $\lambda = h(x)$ and using the fact that h is nonincreasing we have the estimate

$$h(x) \leqslant \|h\|_1/x.$$

Now

$$\int_0^1 x^{-1} \int_0^x h(y)\,dy\,dx = \lim_{x\to 0} \log(1/x) \int_0^x h(y)\,dy + \int_0^1 \log(1/x)\,h(x)\,dx$$

so

$$\int_0^1 x^{-1} \int_0^x h(y)\,dy\,dx \geqslant \int_0^1 \log_+(1/x)\,h(x)\,dx \geqslant \int_0^1 h(x)\log_+\big(h(x)/\|h\|_1\big)\,dx,$$

which yields the left-hand inequality of the lemma.

For the right-hand inequality we may assume $|h|_{L\log_+ L} < \infty$. For $\lambda > 0$ let $E(\lambda) = \{y: \sqrt{y}\,h(y) \geqslant \lambda\}$. Then by separate estimates on and off $E(1)$ we have

$$\log(1/x) \int_0^x h(y)\,dy \leqslant 2\sqrt{x}\log(1/x) + 2\int_0^x h(y)\log_+ h(y)\,dy.$$

Thus, integration by parts yields

$$\int_0^1 x^{-1} \int_0^x h(y)\,dy\,dx = \int_0^1 \log(1/x)\,h(x)\,dx = \int_{E(\lambda)} + \int_{E(\lambda)^c}$$

$$\leqslant 2\int_0^1 h(x)\log_+\big(h(x)/\lambda\big)\,dx + 4\lambda.$$

Taking $\lambda = \|h\|_1/2$ yields the desired inequality.

Remark 3.1. It is now easy to give an example of a function f for which the iterated maximal functions $m_1(m_2(f))$ and $m_2(m_1(f))$ are infinite a.e., and yet $\underline{M}(f) < \infty$ a.e. In fact, the function f of Example 3.1 will work. To see this, fix (x_0, y_0) with $x_0 \neq y_0$ and $0 < |x_0 - y_0| \leqslant 1$. Then

$$m_1(f)(x_0, y_0) \geqslant \frac{1}{|x_0-y_0|} \int_0^{|x_0-y_0|} h(x_0 - x - y_0)\,dx = \frac{1}{|x_0-y_0|} \int_0^{|x_0-y_0|} h(x)\,dx.$$

Since $|h|_{L\log_+ L} = \infty$, Lemma 3.1 implies that $m_1(f)(x_0, y_0)$ is not integrable as a function of y_0. Thus $m_2(m_1(f)) = +\infty$ a.e.

We conclude by showing that dyadic step functions are dense in the quasi-normed function space determined by δ_0. This fact was used above to deduce Corollary 1.1 from Theorem 1.1.

PROPOSITION 3.1. *Let f satisfy $\delta_0(f) < \infty$. Then there exists a sequence f_n of dyadic functions satisfying $\delta_0(f - f_n) \to 0$ and $f_n \to f$ a.e.*

Proof. The following elementary inequalities, valid for $A > 0$ and $B > 0$, are useful in the analysis of δ_0:

$$\log_+(AB) \leqslant \log_+ A + \log_+ B, \tag{3.5}$$

$$A\log_+(B/A) \leqslant B/e \quad (e = 2.718\ldots). \tag{3.6}$$

For example, using (3.5) to split the $\log_+$ expression into 4 terms, performing the obvious integrations and then applying (3.6) yields the estimate

$$\delta_0(f) \leqslant |f|_{L\log_+ L} + (2e^{-1} + \log_+\|f\|_1)\|f\|_1.$$

Since dyadic step functions are dense in the Orlicz spaces L^1 and $L\log_+ L$, the result of the proposition holds whenever f is, say, bounded. Thus, by the quasi-norm property of δ_0, it is enough to prove the following:

Let $g_N(x) = \begin{cases} |f(x)|, & |f(x)| > N, \\ 0, & \text{otherwise}. \end{cases}$ *Then* $\delta_0(g_N) \to 0$.

Now, by (3.5), we have for each $\varepsilon > 0$,

$$\delta_0(g_N) \leqslant \iint g_N(x, y)\left(1 + \log_+ \frac{\|f\|_1\, g_N(x, y)}{\|f(x, \cdot)\|_1\, \|f(\cdot, y)\|_1}\right) dx\,dy$$

$$+ \int_0^1 \|g_N(x, \cdot)\|_1 \log_+\big(\varepsilon\|f(x, \cdot)\|_1/\|g_N(x, \cdot)\|_1\big)\,dx$$

$$+ \int_0^1 \|g_N(\cdot, y)\|_1 \log_+\big(\varepsilon\|f(\cdot, y)\|_1/\|g_N(\cdot, y)\|_1\big)\,dy$$

$$+ \|g_N\|_1 \log_+(\varepsilon^{-2}\|g_N\|_1/\|f\|_1).$$

The desired result follows easily since the first term tends to 0 by the Dominated Convergence Theorem and, by (3.6), the second and third terms are each bounded by $(\varepsilon/e)\|f\|_1$.

Remark 3.2. Similar estimates show that dyadic step functions are dense in the spaces determined by Δ and δ_*.

References

[1] N. A. Fava, E. A. Gatto and C. Gutiérrez, *On the strong maximal function and Zygmund's class* $L(\mathrm{Log}\, L)^n$, Studia Math. 69 (1980), 155–158.

[2] M. de Guzmán, *Differentiation of Integrals in* $\mathbf{R}^n$, Lecture Notes in Math. 481, Springer, Berlin 1975.

[3] G. H. Hardy and J. E. Littlewood, *A maximal theorem with function-theoretic applications*, Acta Math. 54 (1930), 81–116.

[4] B. Jessen, J. Marcinkiewicz and A. Zygmund, *Note on the differentiability of multiple integrals*, Fund. Math. 25 (1935), 217–234.

[5] S. Kwapień and W. A. Woyczyński, *Double stochastic integrals, random quadratic forms and random series in Orlicz spaces*, Ann. Probab. 15 (1987), 1072–1096.
[6] T. R. McConnell, *A two parameter maximal ergodic theorem with dependence*, Ann. Probab., to appear.
[7] —, *Triple stochastic integration with respect to stable processes*, preprint, 1985.
[8] —, *Two-parameter strong laws and maximal inequalities for U-statistics*, Proc. Roy. Soc. Edinburgh, to appear.
[9] T. R. McConnell and M. S. Taqqu, *Dyadic approximation of double integrals with respect to symmetric stable processes*, Stochastic Process. Appl., to appear.
[10] —, —, *Decoupling of Banach-valued multi-linear forms in independent symmetric Banach-valued random variables*, Probab. Related Fields, to appear.
[11] E. M. Stein, *Note on the class* $L\log L$, Studia Math. 32 (1969), 305–310.

DEPARTMENT OF MATHEMATICS
SYRACUSE UNIVERSITY
200 Carnegie, Syracuse, New York 13244-1150, U.S.A.

Received September 2, 1986
Revised version November 24, 1986 (2209)