Finally, we establish the

Sufficiency of (3). From the well-known "openness" of the condition for membership in A_p, w belongs to A_{r_0} for some r_0 with $1<r_0<p$, where p^{-1} is the upper index of L_Φ. Then, for Lebesgue-measurable f,

$$\int_{R^n} \Phi\big((Mf)(x)\big)w(x)\,dx \leqslant C\int_0^\infty w(\{Mf>s\})\Phi(s)\,(ds/s)$$
$$\leqslant C\int_0^\infty (\Phi(s)/s^{r_0})\Big(\int_{|f(x)|>s} |f(x)|^{r_0}w(x)\,dx\Big)(ds/s).$$

Interchanging the order of integration gives

$$\int_{R^n} |f(x)|^{r_0}w(x)\Big(\int_0^{|f(x)|}(\Phi(s)/s^{r_0})(ds/s)\Big)\,dx.$$

But, using (8) with, say, $r=(r_0+p)/2$, means this is dominated by a constant multiple of $\int_{R^n}\Phi\big(|f(x)|\big)w(x)\,dx$. This completes our proof.

References

[1]. D. W. Boyd, *Indices of function spaces and their relationship to interpolation*, Canad. J. Math. 21 (1969), 1245-1254.

[2] R. R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. 54 (1974), 221–237.

[3] W. Matuszewska and W. Orlicz, *On certain properties of φ-functions*, Bull. Acad. Polon. Sci. 8, 7 (1960), 439–443.

[4] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207–226.

[5] — *The equivalence of two conditions for weight functions*, Studia Math. 49 (1974), 101–106.

[6] E. M. Stein and G. Weiss, *An extension of a theorem of Marcinkiewicz and some of its applications*, J. Math. Mech. 8 (1959), 263–264.

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA, 47405
and
DEPARTMENT OF MATHEMATICS
BROCK UNIVERSITY
ST. CATHARINES, ONTARIO
CANADA, L2S3A1
DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

Received June 11, 1979 (1552)

STUDIA MATHEMATICA, T. LXXI. (1982)

On the convergence of bilinear and quadratic forms in independent random variables

by

PETER SJÖGREN (Göteborg)

Abstract. We consider bilinear and quadratic forms $\sum a_{ij}X_iY_j$ and $\sum a_{ij}X_iX_j$ in independent random variables with expectations 0 and variances 1. Necessary and sufficient conditions for these forms to converge a.s. are given. When the X_i and Y_j are normal, we consider $X=(X_i)$ and $Y=(Y_j)$ as vectors in $\boldsymbol{R}^N$ and ask when $\sum a_{ij}X_iY_j$ converges for X and Y in a subspace of $\boldsymbol{R}^N$ of measure 1 for the distribution law. This is proved to happen precisely when the a_{ij} define a nuclear operator on l^2. The natural extension of this theorem to trilinear forms is shown to be false. An analogous result for stochastic integrals is also given.

1. Introduction and statements of results. In this paper all random variables and coefficients will be real-valued. However, our results extend to the complex-valued case with only small modifications. The probability measure is denoted by P.

We shall say that a set of random variables *stays away from* 0 if it contains no sequence tending to 0 in probability, i.e., if there is an $\varepsilon>0$ such that $P(|X|\geqslant\varepsilon)\geqslant\varepsilon$ for all X in the set.

Linear forms $\sum a_iX_i$ have been considered by Hoffmann-Jørgensen [2], Th. 4.10. If the X_i are independent, stay away from 0, and satisfy $EX_i=0$ and $EX_i^2=1$, then the condition $\sum a_i^2<\infty$ is necessary and sufficient for $\sum a_iX_i$ to converge a.s. Notice that this conclusion holds if and only if the X_i stay away from 0, when the other assumptions are satisfied.

For bilinear forms, several kinds of convergence exist. Call (M_k, N_k) an *admissible sequence* if each M_k and N_k is a natural number or ∞, and M_k and N_k increase to ∞ with k, and both M_k and N_k are not ∞. A bilinear form $\sum' a_{ij}X_iY_j$ is said to *converge* for such a sequence if $\sum_{i<M_k, j<N_k} a_{ij}X_iY_j$ converges as $k\to\infty$. Hoffmann-Jørgensen's techniques [2], pp. 155–156, are easily modified to give the following result.

THEOREM 1. *Let X_i and Y_j, $i,j=1,2,\ldots$, be independent, have expectation 0 and variance 1, and stay away from 0. Then the following are equivalent:*

(i) $\sum_{i,j} a_{ij}^2 < \infty$;

(ii) $\sum_{i<M, j<N} a_{ij} X_i Y_j$ *converges* a.s. *as* $M, N \to \infty$;

(iii) $\sum a_{ij} X_i Y_j$ *converges* a.s. *for some admissible sequence.*

As to quadratic forms, we must pay special attention to the diagonal terms since they have nonvanishing expectations.

THEOREM 2. *Let* X_i, $i = 1, 2, \ldots$, *be independent with expectation* 0 *and variance* 1. *Assume further that* $EX_i^4 < \infty$ *for all* i *and that the set*

$$\{(X_i^2-1)/\big(E(X_i^2-1)^2\big)^{1/2}\colon P(|X_i| = 1) < 1\}$$

stays away from 0. *Also let* $a_{ij} = a_{ji}$ *for all* i *and* j. *Then* $\sum_{i,j<N} a_{ij} X_i X_j$ *converges* a.s. *as* $N \to \infty$ *if and only if*

$$\sum_{i \neq j} a_{ij}^2 + \sum a_{ii}^2 E(X_i^2-1)^2 < \infty$$

and $\lim_{N\to\infty} \sum_{i<N} a_{ii}$ *exists in* $\mathbf{R}$.

When the X_i are identically distributed, the variables $(X_i^2-1)/\big(E(X^2-1)^2\big)^{1/2}$ will always stay away from 0, except when $|X_i| = 1$ a.s. We state this special case as a corollary.

COROLLARY. *Let* X_i, $i = 1, 2, \ldots$, *be independent and identically distributed and satisfy* $EX_i = 0$, $EX_i^2 = 1$, *and* $EX_i^4 < \infty$. *Assume* $a_{ij} = a_{ji}$. *If* $P(|X_i| = 1) < 1$, *then* $\sum_{i,j<N} a_{ij} X_i X_j$ *converges* a.s. *as* $N \to \infty$ *if and only if* $\sum_{i,j} a_{ij}^2 < \infty$ *and* $\lim_{N\to\infty} \sum_{i<N} a_{ii}$ *exists in* $\mathbf{R}$. *If* $P(|X_i| = 1) = 1$, *i.e.*, $P(X_i = \pm 1) = 1/2$, *the same conclusion holds provided we replace* $\sum_{i,j} a_{ij}^2$ *by* $\sum_{i\neq j} a_{ij}^2$.

Under slightly stronger assumptions, Theorem 2 follows from Schreiber [4]. It is not hard to see that in each of Theorems 1 and 2 the hypothesis saying that a set stays away from 0 is necessary and sufficient for the conclusion to hold, given the other hypotheses.

We now adopt a more functional-analytic standpoint and consider $X = (X_i)_1^\infty$ and $Y = (Y_j)_1^\infty$ as vectors in $\mathbf{R}^N$. Assume the Y_j are independent copies of the X_i so that X and Y have the same distribution law μ, which is considered as a measure in $\mathbf{R}^N$. If the a_{ij} satisfy the hypotheses of Theorems 1 and 2, we have a bilinear form $B(X, Y)$ and a quadratic form $Q(X)$, and $Q(X) = B(X, X)$ when $X \in D(Q)$, the domain of Q. When the a_{ij} are symmetric, it is natural to ask whether one can reconstruct B from Q, say by means of

$$B(X, Y) = \tfrac{1}{2}\big(Q(X+Y) - Q(X) - Q(Y)\big) \tag{1.1}$$

for $X, Y \in D(Q)$. Then $X+Y$ should also be in $D(Q)$, i.e., $D(Q)$ must be a vector subspace of μ-measure 1. This is equivalent to saying that $B(X, Y)$ is defined when X and Y belong to some subspace of μ-measure 1. As is easily seen, it is equivalent to assume B defined on the product of two subspaces, or sets, of μ-measure 1. This property of B makes sense even if B is not symmetric. For normal variables there is a simple characterization of the coefficients of such forms. By a *full subspace* we mean a μ-measurable subspace of $\mathbf{R}^N$ of μ-measure 1.

THEOREM 3. *Let* X_i *and* Y_j, $i, j = 1, 2, \ldots$, *be independent and* $N(0, 1)$, *so that* μ *is the canonical Gaussian measure in* $\mathbf{R}^N$. *Then the following are equivalent*:

(i) *the matrix* (a_{ij}) *defines a nuclear operator* $A\colon l^2 \to l^2$;

(ii) $\sum_{i<M, j<N} a_{ij} X_i Y_j$ *converges for* X *and* Y *in a full subspace as* $M, N \to \infty$;

(iii) *the same sum converges for* X *and* Y *in a full subspace as* (M, N) *runs through some admissible sequence.*

If $a_{ij} = a_{ji}$, *this is also equivalent to*

(iv) $\sum_{i,j<N} a_{ij} X_i X_j$ *converges on a full subspace as* $N \to \infty$.

This means that convergence on full subspaces is far stronger than the a.s. convergence of Theorem 1. Simple examples of matrices (a_{ij}) having the property of Theorem 1 but not that of Theorem 2 are obtained by letting $a_{ij} = 0$ for $i \neq j$ and taking as (a_{ii}) a sequence in l^2 but not in l^1.

It seems plausible that Theorem 3 is valid for more general distributions. In fact, our proof of the implication (i) $\Rightarrow$ (ii) holds quite generally, but to show (iii) $\Rightarrow$ (i), we use techniques applying only to normal variables.

Next, we study analogous forms for stochastic integrals. Let β and γ be independent Brownian motions and K a sure function in $L^2(R)$, $R = [0, 1] \times [0, 1]$. Then the stochastic integral $\iint_R K(s, t)\, d\beta(s)\, d\gamma(t)$ exists. One way of defining it is to consider partitions Δ: $0 = s_0 < s_1 < \ldots < s_m = 1$ and $0 = t_0 < t_1 < \ldots < t_n = 1$, and form the Riemann sum

$$\iint K_\Delta(s, t)\, d\beta(s)\, d\gamma(t) = \sum_{i,j} m_{ij}\big(\beta(s_i) - \beta(s_{i-1})\big)\big(\gamma(t_j) - \gamma(t_{j-1})\big),$$

where K_Δ is the step function whose value in a rectangle $]s_{i-1}, s_i[\times]t_{j-1}, t_j[$ equals the mean value m_{ij} of K in this rectangle. By an *admissible sequence of partitions* we mean a sequence $(\Delta_k)_1^\infty$ such that Δ_{k+1} is a refinement of Δ_k and the largest rectangle side of Δ_k tends to 0 as $k \to \infty$. For such a sequence, $\iint K_{\Delta_k} d\beta d\gamma$ is an L^2 bounded martingale in k, as we shall see later, and converges a.s. to $\iint K\, d\beta d\gamma$. Let ν be the Wiener measure in $C[0, 1]$. The following result is analogous to Theorem 3.

THEOREM 4. *Let* β, γ, *and* K *be as just described, and assume* $K \in L^2(R)$. *The following are equivalent*:

(i) *the operator* $K: L^2[0,1]\to L^2[0,1]$ *defined by* $K\varphi(s) = \int K(s,t)\times\varphi(t)\,dt$ *is nuclear*;

(ii) *as* Δ *runs through any admissible sequence of partitions,* $\iint K_\Delta\, d\beta\, d\gamma$ *converges for* β *and* γ *in a subspace of* $C[0,1]$ *which is full with respect to* ν;

(iii) *same as* (ii) *but only for some admissible sequence.*

If $K(s,t) = K(t,s)$ *and only symmetric partitions are considered, this is also equivalent to*

(iv) $\iint K_\Delta(s,t)\,d\beta(s)\,d\beta(t)$ *converges for any* (*or for some*) *admissible sequence for* β *in a* ν*-full subspace.*

Notice that the subspaces in (ii)–(iv) depend on the admissible sequence considered. In general, one does not have convergence in the directed set of all partitions, see Sjögren [6].

The proofs of Theorems 1–4 constitute Sections 2–4.

Finally, we consider trilinear forms $\sum a_{ijk}X_iY_jZ_k$ in independent $N(0,1)$ variables X_i, Y_j, Z_k. By the method of proof of Theorem 3, it can be shown that such a form converges on the product of three full subspaces if the coefficient tensor is in the completed projective tensor product $l^2\hat{\otimes}_\pi l^2\hat{\otimes}_\pi l^2$. But the converse of this is false, as proved by means of an example in Section 5.

Our main results were given in the preliminary report Sjögren [5].

By C we denote many different positive constants whose values are unimportant.

2. Almost sure convergence.

LEMMA 1. *Let* X_i, $i = 1,2,\ldots$, *be independent, have expectation* 0 *and variance* 1, *and stay away from* 0. *Then the set* $\{\sum a_iX_i:\ \sum a_i^2 = 1\}$ *stays away from* 0.

Proof. Assume $P(|X_i|\geqslant\varepsilon)\geqslant\varepsilon>0$ for all i. By Chebyshev's inequality, $P(\varepsilon\leqslant|X_i|\leqslant\sqrt{2/\varepsilon})\geqslant\varepsilon/2$, so for small t one has

$$|E\cos tX_i|\leqslant 1-\varepsilon^3t^2/8.$$

Further,

$$|E\sin tX_i| = |E(\sin tX_i - tX_i)|\leqslant E(t^2X_i^2) = t^2.$$

So the characteristic function of X_i satisfies

$$|E\exp(\mathrm{it}X_i)|^2\leqslant(1-\varepsilon^3t^2/8)^2+t^4\leqslant\exp(-\varepsilon^3t^2/8),$$

for $|t|\leqslant t_0 = t_0(\varepsilon)$. For these t then

$$\Big|E\exp\Big(it\sum a_iX_i\Big)\Big|^2\leqslant\prod\exp(-\varepsilon^3t^2a_i^2/8) = \exp(-\varepsilon^3t^2/8).$$

The lemma follows.

LEMMA 2. *If the random variables* Z_i, $i=1,2,\ldots$, *stay away from* 0, *then there exists a* $\delta>0$ *such that for every sequence* $c_i\geqslant 0$, $i=1,2,\ldots$, *one has*

$$P\Big(\sum c_iZ_i^2\geqslant\delta\sum c_i\Big)\geqslant\delta.$$

Proof. Let $U_i = \min(Z_i^2, 1)$, and notice that $EU_i\geqslant\varepsilon$ for some $\varepsilon>0$. If $\sum c_i<\infty$, then $T = \sum c_iU_i/\sum c_i$ is bounded by 1 and has expectation at least ε. Hence, $P(T\geqslant\delta)\geqslant\delta$, $\delta=\delta(\varepsilon)$, from which the assertion follows. The case $\sum c_i = \infty$ is easily handled, and the proof is complete.

Proof of Theorem 1. (i) $\Rightarrow$ (ii). The partial sums in (ii) form a double index martingale which is bounded in L^2 because of (i). From this (ii) follows, in view of Théorème 2, p. 6 in Cairoli [1].

Since (ii) trivially implies (iii), we prove (iii) $\Rightarrow$ (i). Let μ_X and μ_Y be the distribution laws of X and Y, respectively. By Lemma 1, there is an $\varepsilon_1>0$ such that for any fixed Y and any M and N

$$\mu_X\Big\{X:\ \Big(\sum_{i<M}X_i\sum_{j<N}a_{ij}Y_j\Big)^2\geqslant\varepsilon_1\sum_{i<M}\Big(\sum_{j<N}a_{ij}Y_j\Big)^2\Big\}\geqslant\varepsilon_1. \tag{2.1}$$

Normalizing, we let

$$\Big(\sum_{j<N}a_{ij}Y_j\Big)^2 = \Big(\sum_{j<N}a_{ij}^2\Big)\Big(\sum_{j<N}\tilde a_{ij}Y_j\Big)^2.$$

Then Lemmas 1 and 2 imply

$$\mu_Y\Big\{Y:\ \sum_{i<M}\Big(\sum_{j<N}a_{ij}Y_j\Big)^2\geqslant\varepsilon_2\sum_{i<M,j<N}a_{ij}^2\Big\}\geqslant\varepsilon_2, \tag{2.2}$$

for some $\varepsilon_2>0$. Since X and Y are independent, (2.1), (2.2) show that

$$P\Big(\Big(\sum_{i<M,j<N}a_{ij}X_iY_j\Big)^2\geqslant\varepsilon_1\varepsilon_2\sum_{i<M,j<N}a_{ij}^2\Big)\geqslant\varepsilon_1\varepsilon_2.$$

Thus, if $\sum a_{ij}X_iY_j$ converges a.s. for some admissible sequence, $\sum a_{ij}^2<\infty$. This ends the proof of Theorem 1.

Proof of Theorem 2. The hypotheses clearly imply

$$E(|X_i^2-1|)\geqslant\varepsilon\big(E(X_i^2-1)^2\big)^{1/2},$$

for some $\varepsilon>0$. This gives an estimate for $E(X_i^2-1)^2$ from which it follows that

$$EX_i^4\leqslant C. \tag{2.3}$$

Take a natural number N. In the rest of this section, it is understood that $1\leqslant i,j\leqslant N$ in all sums. Let $Q_N = \sum a_{ij}X_iX_j = R_N+T_N$, where $R_N = \sum a_{ii}(X_i^2-1)$, so that $T_N = \sum_{i\neq j}a_{ij}X_iX_j+m_N$, with $m_N = \sum a_{ii}$.

Then R_N and T_N are orthogonal in L^2 and have norms

$$r_N = \|R_N\|_2 = \Big(\sum a_{ii}^2 E(X_i^2-1)^2\Big)^{1/2} \quad \text{and} \quad t_N = \|T_N\|_2 = \Big(4\sum_{j<i} a_{ij}^2 + m_N^2\Big)^{1/2}.$$

Also, $\|Q_N\|_2 = (r_N^2+t_N^2)^{1/2}$.

LEMMA 3. *The L^4 norm of T_N satisfies $\|T_N\|_4 \leqslant Ct_N$.*

Proof. Since $\|T_N\|_4 \leqslant \|\sum_{i\neq j} a_{ij}X_iX_j\|_4 + |m_N|$, it is enough to prove

$$E\Big(\Big(\sum_{j<i} a_{ij}X_iX_j\Big)^4\Big) = O(t_N^4). \tag{2.4}$$

Developing the fourth power in (2.4), we obtain terms of type $\prod_{k=1}^4 a_{i_kj_k}X_{i_k}X_{j_k}$ with $j_k < i_k$. As soon as some subscript i_k or j_k appears only once in such a product, the expectation of this product vanishes. We classify the remaining terms according to the number n of distinct pairs (i_k, j_k) it contains. For $n = 1$, we get terms $a_{ij}^4X_i^4X_j^4$, and because of (2.3) they have a total expectation bounded by $C\sum_{j<i} a_{ij}^4 = O(t_N^4)$. When $n = 2$, we get terms $a_{ij}^2a_{kl}^2X_i^2X_j^2X_k^2X_l^2$. The expectation of such an expression is $O(a_{ij}^2a_{kl}^2)$, because of (2.3) and since at most two of these four subscripts may coincide. Again the sum of the expectations is $O(t_N^4)$. For $n = 3$, the terms are necessarily of type $a_{ij}^2a_{ik}a_{jk}X_i^3X_j^3X_k^2$ (except possibly for the order between the subscripts of each a). The expectation of this is at most $Ca_{ij}^2a_{ik}a_{jk} \leqslant Ca_{ij}^2a_{ik}^2 + Ca_{ij}^2a_{jk}^2$, and we can sum as before. Finally, $n = 4$ gives

$$\sum E(a_{ij}a_{jk}a_{kl}a_{li}X_i^2X_j^2X_k^2X_l^2) \leqslant \sum a_{ij}^2a_{kl}^2 + a_{jk}^2a_{li}^2 = O(t_N^4),$$

which completes the proof of (2.4) and the lemma.

LEMMA 4. *The random variables $Q_N/\|Q_N\|_2$, $\|Q_N\|_2 \neq 0$, stay away from 0.*

Proof. Because of Lemma 1 and the hypotheses of Theorem 2, the variables R_N/r_N stay away from 0. If $t_N/r_N \leqslant c_0$ for some small $c_0 > 0$ which can be taken independent of N, it is easy to see that also the variables Q_N/r_N and $Q_N/(r_N^2+t_N^2)^{1/2}$ stay away from 0.

Consider next those N for which $t_N/r_N > c_0$. Let A be the event $\{|Q_N| > \delta(r_N^2+t_N^2)^{1/2}\}$ and assume $P(A) < \delta$, $\delta > 0$. Then by Minkowski's inequality and Lemma 3,

$$\Big(\int_A Q_N^2\,dP\Big)^{1/2} \leqslant \Big(\int_A R_N^2\,dP\Big)^{1/2} + \Big(\int_A T_N^2\,dP\Big)^{1/2}$$
$$\leqslant r_N + P(A)^{1/4}\Big(\int_A T_N^4\,dP\Big)^{1/4} \leqslant r_N + C\delta^{1/4}t_N.$$

Estimating $\int_A Q_N^2\,dP$ from below, we get

$$(1-\delta^2)(r_N^2+t_N^2) \leqslant (r_N + C\delta^{1/4}t_N)^2.$$

But this is impossible when $t_N/r_N > c_0$, if δ is small enough, as is easily verified. Lemma 4 is proved.

End of proof of Theorem 2. The "if" part follows from the fact that $(Q_N - m_N)_{N=1}^\infty$ is an L^2 bounded martingale.

Conversely, if Q_N converges a.s., there is for any $\varepsilon > 0$ an M such that $P(|Q_N| \leqslant M$, all $N) > 1-\varepsilon$. From Lemma 4, we see that r_N and t_N must be bounded. But then $(Q_N - m_N)_1^\infty$ is again convergent a.s., and so m_N converges. This ends the proof of Theorem 2.

3. Convergence on subspaces. Let A be a compact linear operator in l^2. Then A^*A has a spectral decomposition $A^*Ax = \sum\lambda_p\langle x, e_p\rangle e_p$, where $\lambda_p > 0$ and the e_p form an orthonormal system. Set $c_p = \sqrt{\lambda_p}$ and $f_p = c_p^{-1}Ae_p$, which defines another orthonormal system. This gives a diagonalization $Ax = \sum c_p\langle x, e_p\rangle f_p$ of A with $c_p > 0$. Then A is nuclear if and only if $\sum c_p < \infty$, and $\|A\|_{\text{nucl}} = \sum c_p$.

LEMMA 5. *Let $Ax = \sum c_p\langle x, e_p\rangle f_p$ and $A'x = \sum c_p'\langle x, e_p'\rangle f_p'$ be two compact operators with diagonalizations as just described. Assume the c_p and c_p' are numbered in decreasing order. Then $|c_p - c_p'| \leqslant \|A - A'\|$ for every p, where we use the operator norm.*

Proof. This follows immediately from the simple minimax formula for c_p which says that

$$c_p = \inf\{\eta(L)\colon L \subset l^2 \text{ subspace of dimension } p-1\},$$

where

$$\eta(L) = \sup\{\|Ax\|\colon x \in L^\perp \text{ and } \|x\| \leqslant 1\}.$$

Proof of Theorem 3. (i)$\Rightarrow$(ii). Let A be the nuclear operator defined by (a_{ij}), and put $X^M = (X_1, \ldots, X_{M-1}, 0, \ldots)$ and similarly for Y^N. Then

$$S_{M,N} = \sum_{i<M, j<N} a_{ij}X_iY_j = \langle X^M, AY^N\rangle$$

for M and N finite, and thus

$$S_{M,N} = \sum c_p\langle X^M, f_p\rangle\langle Y^N, e_p\rangle \tag{3.1}$$

if we diagonalize A as before. Now observe that $\langle X^M, f_p\rangle_{M=1}^\infty$ is for each p a martingale whose L^2 norm is 1. As $M\to\infty$ it converges a.s. to $\langle X, f_p\rangle$. By the L^2 maximal theorem for martingales, we have

$$\|\sup_M|\langle X^M, f_p\rangle|\|_2 \leqslant 2. \tag{3.2}$$

Further,

$$|c_p\langle X^M, f_p\rangle\langle Y^N, e_p\rangle| \leqslant \tfrac{1}{2}c_p\langle X^M, f_p\rangle^2 + \tfrac{1}{2}c_p\langle Y^N, e_p\rangle^2 \leqslant \tfrac{1}{2}c_p\sup_M\langle X^M, f_p\rangle^2 + \tfrac{1}{2}c_p\sup_N\langle Y^N, e_p\rangle^2. \tag{3.3}$$

Set

$$F_X = \Big\{X\colon \sum_p c_p\sup_M\langle X^M, f_p\rangle^2 < \infty \text{ and } \langle X_M, f_p\rangle_{M=1}^\infty \text{ converges for each } p\Big\},$$

which is a subspace of $\boldsymbol{R}^N$ and of full measure because of (3.2) and the fact that $\sum c_p < \infty$. Defining F_Y analogously, we see that for $(X, Y) \in F_X \times F_Y$ we may apply (3.3) to deduce the convergence of $S_{M,N}$ as $M, N \to \infty$.

(ii) $\Rightarrow$(iii). Trivial.

(iii) $\Rightarrow$(i). Because of Theorem 1, it is enough to assume A Hilbert–Schmidt but not nuclear, and prove that we cannot have convergence on a product of full subspaces for a given admissible sequence (M_k, N_k). Suppose first that all M_k and N_k are finite, and put $K_k = \min(M_k, N_k)$ and $L_k = \max(M_k, N_k)$.

For some k, let $A^1\colon l^2 \to l^2$ be the truncation of A in $i < M_k$, $j < N_k$, i.e., the operator defined by the matrix whose entries are a_{ij} for $i < M_k$, $j < N_k$ and 0 otherwise. From Lemma 5, we see that $\|A^1\|_{\mathrm{nucl}} > 1$ if k is large enough, and picking a subsequence of the given admissible sequence, we may assume $k = 1$. As before, we may diagonalize A^1, getting

$$\langle X, A^1 Y\rangle = \sum_{p<K_1} c_p X'_p Y'_p,$$

where the X'_p are orthonormal linear combinations of X_i, $i < M_k$, and similarly for the Y'_p.

Let A_1 be the truncation of A in $i \geqslant L_1$, $j \geqslant L_1$. Since $A - A_1$ is of finite rank and hence nuclear, A_1 is not nuclear. Call A_1^2 the *truncation of A in* $L_1 \leqslant i < M_2$, $L_1 \leqslant j < N_2$. Passing again to a subsequence, we may assume that M_2 and N_2 are so large that $\|A_1^2\|_{\mathrm{nucl}} > 4\|A - A_1\|_{\mathrm{nucl}} + 2$. Then diagonalize A_1^2 so that

$$\langle X, A_1^2 Y\rangle = \sum_{L_1 \leqslant p < K_2} c_p X'_p Y'_p.$$

This time, the X'_p are orthogonal linear combinations of X_i, $L_1 \leqslant i < M_2$. Continuing in this way, we let A_{k-1} be the truncation of A in $i \geqslant L_{k-1}$, $j \geqslant L_{k-1}$, and A_{k-1}^k the truncation of A in $L_{k-1} \leqslant i < M_k$, $L_{k-1} \leqslant j < N_k$, and assume $\|A_{k-1}^k\|_{\mathrm{nucl}} > 4\|A - A_{k-1}\|_{\mathrm{nucl}} + k$. Then diagonalize A_{k-1}^k so that

$$\langle X, A_{k-1}^k Y\rangle = \sum_{L_{k-1} \leqslant p < K_k} c_p X'_p Y'_p.$$

We have

$$S_k = S_{M_k, N_k} = \sum_{L_{k-1} \leqslant p < K_k} c_p X'_p Y'_p + \langle X^{M_k}, (A - A_{k-1}) Y^{N_k}\rangle. \tag{3.4}$$

For $K_k \leqslant p < L_k$, we define X'_p in a suitable way so as to obtain a complete system $X' = (X'_p)_1^\infty$ of orthonormal linear combinations of the X_i, and similarly for Y'_p.

If S_k converges for X and Y in a full subspace, then it converges for $X' \in F_X$, $Y' \in F_Y$ for full subspaces F_X and F_Y, since orthogonal transformations are μ-invariant. Thus we have convergence for $X' = Y' \in F_X \cap F_Y$. Consider for a fixed k the bilinear form $\langle X^{M_k}, (A - A_{k-1}) Y^{N_k}\rangle$ as a function $B_1(X', Y')$ of (X', Y'). Diagonalizing and using the estimates of the proof of (i) $\Rightarrow$(ii) we deduce

$$\int |B_1(X', X')|\, d\mu(X') \leqslant 2\|A - A_{k-1}\|_{\mathrm{nucl}}. \tag{3.5}$$

One can clearly choose $a < 1$ such that if S is an event with $P(S) > a$, then $\int_S Z^2 dP > 1/2$ for any $N(0, 1)$ variable Z. Thus, for any μ-measurable set $E \subset \boldsymbol{R}^N$ with $\mu(E) > a$, one has

$$\int_E \sum_{L_{k-1} \leqslant p < K_k} c_p (X'_p)^2 d\mu > \tfrac{1}{2} \sum_{L_{k-1} \leqslant p < K_k} c_p > 2\|A - A_{k-1}\|_{\mathrm{nucl}} + k/2.$$

From (3.4), (3.5) we then get

$$\int_{X' = Y' \in E} S_k\, d\mu(X') > k/2.$$

Since k is arbitrary, this is not compatible with a.s. convergence for $X' = Y'$, and we are done.

Suppose next $N_k = \infty$ for all k. For operators in l^2 of fixed finite rank, it is easy to estimate the nuclear norm in terms of the operator norm. Therefore, we can find an increasing sequence N'_k such that the nuclear norm of the truncation A'_k of A in $i < M_k$, $j \geqslant N'_k$ stays bounded as $k \to \infty$. Let S'_k denote the partial sums corresponding to the admissible sequence (M_k, N'_k). Then

$$S_k = S'_k + \langle X, A'_k Y\rangle.$$

The nuclear norm of the operator corresponding to S'_k tends to infinity with k. We can now diagonalize S'_k and proceed essentially as we did from (3.4), proving that S'_k dominates $\langle X, A'_k Y\rangle$.

Now assume A symmetric. Clearly (ii) implies (iv), and (iv) implies (iii) with $M_k = N_k = k$, bacause of (1.1).

Theorem 3 is proved.

4. Stochastic integrals.

Proof of Theorem 4. Let $(\Delta_k)_1^\infty$ be an admissible sequence of partitions and assume $]s_{i-1}^k, s_i^k[\times]t_{j-1}^k, t_j^k[$ are the rectangles of Δ_k. Call β_k' the (random) step function which equals $\big(\beta(s_i^k) - \beta(s_{i-1}^k)\big)/(s_i^k - s_{i-1}^k)$ in $]s_{i-1}^k, s_i^k[$. Define γ_k' similarly by means of the t_j^k. Then

$$\int K_{\Delta_k} d\beta\, d\gamma = \int K(s, t)\beta_k'(s)\gamma_k'(t)\, ds\, dt = \langle \beta_k', \boldsymbol{K}\gamma_k' \rangle.$$

Let S_k be the subspace of $L^2 = L^2[0, 1]$ consisting of those step functions which are constant in the intervals $]s_{i-1}^k, s_i^k[$, and set $M_k = \dim S_k$. Then choose an orthonormal basis $f_1, f_2, \ldots$ of L^2 such that $f_1, \ldots, f_{M_k}$ is a basis of S_k, for each k. Choose further a similar basis $e_1, e_2, \ldots$, but with the s_i^k replaced by the t_j^k and M_k by N_k. Then

$$\langle \beta_k', \boldsymbol{K}\gamma_k' \rangle = \sum_{i \leqslant M_k, j \leqslant N_k} a_{ij} \langle \beta_k', f_i \rangle \langle \gamma_k', e_j \rangle,$$

where $a_{ij} = \langle f_i, \boldsymbol{K} e_j \rangle$. Now $\langle \beta_k', f_i \rangle = \int f_i\, d\beta$ for $i \leqslant M_k$, so this scalar product is independent of k and will be called X_i. Clearly X_i is $N(0, 1)$, and since the f_i are orthogonal, the X_i are independent. Similarly, let $Y_j = \langle \gamma_k', e_j \rangle$ for $j \leqslant N_k$. Evidently, $\boldsymbol{K}$ is nuclear if and only if the a_{ij} define a nuclear operator in l^2. Summarizing, we have

$$\int K_{\Delta_k} d\beta\, d\gamma = \sum_{i \leqslant M_k, j \leqslant N_k} a_{ij} X_i Y_j, \tag{4.1}$$

and the equivalence of (i), (ii), and (iii) follows from Theorem 3. As to (iv), cf. the end of Section 3. We also see that the integral in (4.1) defines a martingale in k.

5. A counterexample for trilinear forms. Let $(X_i)_1^\infty$, $(Y_j)_1^\infty$, and $(Z_k)_1^\infty$ be independent $N(0, 1)$ variables. We shall construct a trilinear form $\sum a_{ijk} X_i Y_j Z_k$ which converges on a product of three full subspaces but whose coefficient tensor is not in $l^2 \hat{\otimes}_\pi l^2 \hat{\otimes}_\pi l^2$.

Choose a sequence $(\varrho_n)_1^\infty$ of integers such that $\varrho_1 = 0$ and $\varrho_{n+1} > \varrho_n + n^2 + n$ for $n \geqslant 1$. Within the "cubic block" $\varrho_n < i, j, k < \varrho_{n+1}$ we relabel the variables, setting $X_i^n = X_{\varrho_n + i}$ and $Y_j^n = Y_{\varrho_n + j}$ and $Z_{ij}^n = Z_{\varrho_n + ni + j}$ for $1 \leqslant i, j \leqslant n$. Denote by X^n and Y^n the n-vectors $(X_i^n)_{i=1}^n$ and $(Y_j^n)_{j=1}^n$, respectively, and by Z^n the $n \times n$ matrix $(Z_{ij}^n)_{i,j=1}^n$. Notice that Z^n is a random matrix all of whose entries are independent and $N(0, 1)$. Let l_n^2 be $\boldsymbol{R}^n$ with the l^2 norm, and consider Z^n as an operator on l_n^2 with norm $\|Z^n\|$.

By the method of Mantero and Tonge [3], Proof of Theorem 1.1, it can be proved that $E\|Z^n\|^2 = O(n)$ as $n \to \infty$, where 1 is the best possible exponent. However, we only need the following simpler estimate.

Lemma 6. $E\|Z^n\|^2 = O(n^{3/2})$, $n \to \infty$.

Proof. Comparing the operator norm to the Hilbert–Schmidt norm, we get

$$\begin{aligned} E\|Z^n\|^2 &= E\|(Z^n)^* Z^n\| \leqslant E\|(Z^n)^* Z^n\|_{\mathrm{HS}} \\ &\leqslant \big(E\|(Z^n)^* Z^n\|_{\mathrm{HS}}^2\big)^{1/2} = \Big(E \sum_{i,j,k,l} Z_{ik}^n Z_{jk}^n Z_{il}^n Z_{jl}^n\Big)^{1/2}. \end{aligned}$$

In the last expression, we see that a term in the sum has vanishing expectation unless $i = j$ or $k = l$. The number of remaining terms is therefore $O(n^3)$ from which the lemma follows.

Consider $T = \sum a_n S^n$, where $S^n = \sum_{1 \leqslant i,j \leqslant n} Z_{ij}^n X_i^n Y_j^n$ and $a_n > 0$. Then

$$|S_n| \leqslant \|Z^n\|\,\|X^n\|\,\|Y^n\| \leqslant n^{1/4}\|Z^n\|^2 + n^{-1/4}\|X^n\|^4 + n^{-1/4}\|Y^n\|^4,$$

and so T converges if

$$\sum a_n n^{1/4} \|Z^n\|^2 < \infty$$

and

$$\sum a_n n^{-1/4} \|X^n\|^4 < \infty$$

and similarly for Y^n. Of course, $E\|X^n\|^4 = O(n^2)$. From this and Lemma 6, we see that if

$$\sum a_n n^{7/4} < \infty, \tag{5.1}$$

these inequalities define full subspaces for Z, X, and Y, respectively.

To see that the coefficient tensor of T need not be in $l^2 \hat{\otimes}_\pi l^2 \hat{\otimes}_\pi l^2$, we show that this tensor need not even be in the larger space $V = l^2 \hat{\otimes}_\pi (l^2 \hat{\otimes}_{\mathrm{HS}} l^2)$. Let $(e_i)_1^\infty$ be an orthonormal basis of l^2, and define e_i^n and e_{ij}^n like the X_i^n etc., and consider the tensor $t^n = \sum_{1 \leqslant i,j \leqslant n} e_{ij}^n \otimes e_i^n \otimes e_j^n$. Here $e_i^n \otimes e_j^n$, $1 \leqslant i, j \leqslant n$, form an orthonormal system in $l^2 \hat{\otimes}_{\mathrm{HS}} l^2$, so t^n has diagonal form, and its V norm is n^2. The e_i involved for different values of n are always orthogonal, so the V norm of the coefficient tensor of T is $\sum a_n n^2$. Determining the a_n so that this last series diverges but (5.1) holds, we obtain the desired example.

The tensor considered is not symmetric in (i, j, k), but its symmetrization gives a symmetric example of the same kind.

References

[1] R. Cairoli, *Une inégalité pour martingales à indices multiples et ses applications*, *Séminaire de probabilité IV*, Lecture Notes in Mathematics 124, 1–27, Springer-Verlag, 1970.

[2] J. Hoffmann-Jørgensen, *Integrability of seminorms, the 0-1 law and the affine kernel for product measures*, Studia Math. 61 (1977), 137–159.

[3] A.M. Mantero & A. Tonge, *The Schur multiplication in tensor algebras*, ibid. 68 (1980), 1–24.

[4] M. Schreiber, *Fermeture en probabilité des chaos de Wiener*, C.R. Acad. Sci. Paris, Sér. A 265 (1967), 859–861.

[5] P. Sjögren, *Almost sure convergence of bilinear and quadratic forms in independent random variables*, Uppsala University, Department of Mathematics, Report 13 (1977).

[6] — *Convergence of Riemann sums for stochastic integrals*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 56 (1981), 181–193.

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GÖTEBORG
S-412 96 GÖTEBORG, SWEDEN

Received February 12, 1979 (1514)
Revised version October 8, 1979

STUDIA MATHEMATICA, T. LXXI. (1982)

Colacunary sequences in L-spaces

by

D. J. ALDOUS (Cambridge) and D.H. FREMLIN (Colchester)

Abstract. A sequence $\langle x_n\rangle_{n\in\mathbf{N}}$ in a normed space is 2-colacunary if there is a continuous linear map from $\operatorname{lin}\{x_n\}$ to l^2 taking each x_n to the nth basic unit vector of l^2. Our main result is that any sequence in an L-space which is bounded but not totally bounded has a 2-colacunary subsequence.

1. DEFINITION. Let E be a normed space. A sequence $\langle x_n\rangle_{n\in\mathbf{N}}$ in E is 2-*colacunary* if there is a $\delta > 0$ such that

$$\Big\|\sum_{i\leq n} a_i x_i\Big\| \geqslant \delta\Big(\sum_{i\leq n} |a_i|^2\Big)^{1/2}$$

for any finite sequence $a_0, \ldots, a_n$ of scalars.

For "lacunary" sequences see [7]. Of course one can define "p-colacunary" sequences similarly for any $p \in \,]0, \infty]$; and the results of §§2 and 9 below will generalise. But the central results of this paper (§§4–6) are special to $p = 2$.

2. LEMMA. *Let E be a normed space, $\langle x_n\rangle_{n\in\mathbf{N}}$ a sequence in E.*

(a) *If $\langle x_n\rangle_{n\in\mathbf{N}}$ is 2-colacunary and $x \in E$, there is an $m \in \mathbf{N}$ such that $\langle x_n + x\rangle_{n\geqslant m}$ is 2-colacunary.*

(b) *If $\langle x_n\rangle_{n\in\mathbf{N}}$ is 2-colacunary and $\langle y_n\rangle_{n\in\mathbf{N}}$ is another sequence in E such that $\sum_{n\in\mathbf{N}} \|x_n - y_n\| < \infty$, then $\langle y_n\rangle_{n\geqslant m}$ is 2-colacunary for some $m \in \mathbf{N}$.*

(c) *If F is another normed space and $T\colon E\to F$ a continuous linear operator such that $\langle Tx_n\rangle_{n\in\mathbf{N}}$ is 2-colacunary, then $\langle x_n\rangle_{n\in\mathbf{N}}$ is 2-colacunary.*

Proof. (a) Let $\sigma > 0$ be such that

$$\Big\|\sum_{i\leq n} a_i x_i\Big\| \geqslant 2\sigma\Big(\sum_{i\leq n} |a_i|^2\Big)^{1/2} \quad \forall\, a_0, \ldots, a_n.$$

Suppose, if possible, that $\langle x_n + x\rangle_{n\geqslant m}$ is never 2-colacunary. Then in particular $\langle x_n + x\rangle_{n\in\mathbf{N}}$ is not 2-colacunary, so there are $a_0, \ldots, a_k$ such that

$$\Big\|\sum_{i\leq k} a_i(x_i + x)\Big\| < \sigma\Big(\sum_{i\leq k} |a_i|^2\Big)^{1/2}.$$