

P-UNIFORM CONVERGENCE AND A VECTOR-VALUED STRONG LAW OF LARGE NUMBERS⁽¹⁾

BY
ANATOLE BECK AND DANIEL P. GIESY

Introduction. In recent years there have been a number of theorems patterned on the strong law of large numbers for random variables which take their values in a normed linear space. For an exposition of some such theorems, the reader is referred to [1].

These theorems all take the following form:

PROPOSITION. *Let \mathfrak{X} be a normed linear space with property P and let R be a restriction on sequences of distributions on \mathfrak{X} . Then every independent sequence $\{X_n\}$ of \mathfrak{X} -random variables with expectations 0 and whose sequence of distributions satisfies R also satisfies the strong law of large numbers; i.e.,*

$$n^{-1}(X_1 + \cdots + X_n) \rightarrow 0$$

as $n \rightarrow \infty$ in the norm topology of \mathfrak{X} almost surely.

Past efforts have concentrated on finding pairs of conditions P and R for which the above proposition is true, frequently starting with a restriction R generalized from classical probability theory and finding a P to match. For most of §III of this paper, we take a different viewpoint, proving theorems of the form:

PROPOSITION. *If for a given normed linear space the previous proposition is true for restriction R , then it is also true for restriction R' .*

As an application, we derive a result (Theorem III.13) of the form of the first proposition above which is valid for all normed linear spaces and in which R restricts only the variances of the random variables. This theorem is best possible in the sense indicated there.

Theorem III.5 depends on the fact that in certain families of independent sequences of random variables such that each sequence averages to 0 almost surely the speed of convergence is uniform. §I is devoted to introducing the necessary concepts for sequences of functions and deriving uniform speed of convergence

Received by the editors June 20, 1969.

⁽¹⁾ The research for and preparation of this paper were supported in part by the National Science Foundation under grant number GP-4588 (Beck) and by the Office of Naval Research under contract number Nonr 228(30) and the University of Southern California Research and Publication Fund under an NDEA grant (Giesy).

Copyright © 1970, American Mathematical Society

criteria. Since the results may be of some interest in themselves, we develop them in greater generality than we need for subsequent sections.

Our basic theorem in §I, Theorem I.3(C), states that if F is a family of sequences of functions, and each element of F converges (by which we mean "is T -summable almost everywhere or in measure or locally in measure", see Definition I.1) to 0, if F is closed with respect to "scrambling", a technique of forming new sequences from old, or if every sequence "scrambled from F " converges to 0, and if a certain technical hypothesis is satisfied, then the rate of convergence of elements of F to 0 is uniform.

In §II, we apply the results of §I to independent sequences of vector-valued random variables and give an application (II.8) which extends a theorem of [4].

I. μ -uniform convergence. If (S, Σ, μ) is a σ -finite measure space and f_n and f are measurable functions on S , then equivalent to the usual definition that $f_n \rightarrow f$ μ -a.e. is

(i) for every $A \in \Sigma$ with $\mu(A) < \infty$ and every $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu \left(\bigcup_{k=n}^{\infty} \{x \in A : |f_k(x) - f(x)| \geq \varepsilon\} \right) = 0;$$

while $f_n \rightarrow f$ in measure if

(ii) for every $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \mu\{x \in S : |f_n(x) - f(x)| \geq \varepsilon\} = 0$;

and $f_n \rightarrow f$ locally in measure if

(iii) for every $A \in \Sigma$ with $\mu(A) < \infty$ and every $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu\{x \in A : |f_n(x) - f(x)| \geq \varepsilon\} = 0.$$

Then (i) \Rightarrow (iii) and (ii) \Rightarrow (iii) and if $\mu(S) < \infty$, all reference to arbitrary $A \in \Sigma$ of finite measure can be replaced by S alone without changing (i) or (iii) and (i) \Rightarrow (ii) \Leftrightarrow (iii). These definitions are so formulated that they can be used to define modes of uniform convergence for families of sequences of functions. Emanuel Parzen does this in [6] where he uses such convergence to prove uniform convergence in mean theorems and in [7] where he gives applications to probability theory. In this section we give the necessary definitions and give some theorems stating conditions under which families of sequences converge in one of these uniform modes. Anticipating future needs, all functions will have their range in an arbitrary normed linear space.

I.1. DEFINITIONS. If X is a topological space, then $\mathcal{B}(X)$, the family of Borel sets of X , is the smallest σ -algebra containing all open sets of X . If (S, Σ, μ) is a σ -finite measure space, $f: S \rightarrow X$ is measurable if $f^{-1}(B) \in \Sigma$ for each $B \in \mathcal{B}(X)$ and there exists $N \in \Sigma$ with $\mu(N) = 0$ such that $f(S \setminus N)$ is separable. If Q is an index set and, for each $q \in Q$ and $n \geq 0$, \mathfrak{X}_q is a normed linear space and $f_n(\cdot, q): S \rightarrow \mathfrak{X}_q$ is measurable, then $f_n \rightarrow f_0$

(I) μ -uniformly almost everywhere if for each $A \in \Sigma$ with $\mu(A) < \infty$ and each $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu \left(\bigcup_{k=n}^{\infty} \{x \in A : \|f_k(x, q) - f_0(x, q)\| \geq \varepsilon\} \right) = 0$$

uniformly for $q \in Q$;

(II) μ -uniformly in measure if for each $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu \{x \in S : \|f_n(x, q) - f_0(x, q)\| \geq \varepsilon\} = 0$$

uniformly for $q \in Q$; and

(III) μ -uniformly locally in measure if for each $A \in \Sigma$ with $\mu(A) < \infty$ and each $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu \{x \in A : \|f_n(x, q) - f_0(x, q)\| \geq \varepsilon\} = 0$$

uniformly for $q \in Q$.

As before, (I) \Rightarrow (III) and (II) \Rightarrow (III) and if $\mu(S) < \infty$, the arbitrary A of finite measure in (I) and (III) can be replaced by S and (I) \Rightarrow (II) \Leftrightarrow (III).

Let $T = (a_{ij})$ be a row finite (for each i , $a_{ij} = 0$ for all but finitely many j) doubly infinite matrix of real (or complex, if it makes sense) numbers and \mathfrak{X} a normed linear space. Then T defines a transformation (also denoted by T) mapping sequences from \mathfrak{X} to sequences from \mathfrak{X} as follows. If $x = \{x_n\}$ is a sequence from \mathfrak{X} , $Tx = y$ where $y_n = \sum_{j=1}^{\infty} a_{nj}x_j$. We use the notation $y_n = T_n x$. If in addition, T satisfies

- (a) $\sup_i \sum_{j=1}^{\infty} |a_{ij}| = M < \infty$,
- (b) for each j , $\lim_{i \rightarrow \infty} a_{ij} = 0$, and
- (c) $\lim_{i \rightarrow \infty} \sum_{j=1}^{\infty} a_{ij} = 1$

then T is called a regular method of summability and T preserves limits: If $x = \{x_n\}$ is a sequence in \mathfrak{X} , $\lim_{n \rightarrow \infty} x_n = x_0$, and $y = Tx$, then $\lim_{n \rightarrow \infty} y_n = x_0$. The conditions (a), (b), and (c) are the Silverman-Toeplitz conditions and the results claimed are well-known if \mathfrak{X} is the real or complex numbers [2, II.4.36] even without row finiteness; the extension to arbitrary \mathfrak{X} is routine. Row finiteness is necessary for the domain of T to include all sequences from \mathfrak{X} . In what follows we shall need the even more restrictive condition that T be lower triangular: $a_{ij} = 0$ for all $j > i$.

If T is a regular method of summability on \mathfrak{X} and x is a sequence from \mathfrak{X} such that Tx converges to x_0 , we say x is T -summable to x_0 .

If f is a sequence of \mathfrak{X} -valued, μ -measurable functions and T is a regular method of summability, then the composite Tf is a sequence of \mathfrak{X} -valued, μ -measurable functions. We say f is T -summable to a function g μ -a.e., in measure, or locally in measure if the sequence Tf converges to g in the respective mode. A family F of sequences of functions is T -summable μ -uniformly in one of the three modes if the family $TF = \{Tf : f \in F\}$ is μ -uniformly convergent in the appropriate mode.

If $x = \{x_n\}$ is a sequence from \mathfrak{X} , $P_n x$ is the sequence $\{x_1, \dots, x_n, 0, 0, \dots\}$.

Let F be a family of sequences. A sequence g is a *sequence scrambled from F* if for each n there exists $f \in F$ such that $g_n = f_n$. F is *closed under scrambling* if each

sequence scrambled from F is in F . The *scrambled closure* of F is the set of all sequences scrambled from F ; clearly, the scrambled closure of F is closed under scrambling.

I.2. REMARKS AND EXAMPLES. If T is a regular method of summability and f is a sequence of μ -measurable \mathfrak{X} -valued functions such that f converges to g μ -a.e., then so does Tf ; the situation for convergence in measure is quite different, however, even if S has finite measure.

EXAMPLE 1. Let C be Cesaro $(C, 1)$ summability: $a_{mn} = 1/m$ if $n \leq m$, 0 otherwise. On $[0, 1]$ with Lebesgue measure, let $f_1 = \chi_{[0,1]}$ (χ_A is the characteristic function of A), $f_2 = 2\chi_{[0,1/2]}$, $f_3 = 2\chi_{[1/2,1]}$, $f_4 = 3\chi_{[0,1/3]}$, $f_5 = 3\chi_{[1/3,2/3]}$, $f_6 = 3\chi_{[2/3,1]}$, and so forth. Then f converges to 0 in measure, but Cf contains the function $\chi_{[0,1]}$ infinitely often ($C_n f = \chi_{[0,1]}$ for $n = k(k+1)/2$) so does not converge to 0 in measure. In fact Cf converges to $\chi_{[0,1]}$ uniformly, hence a.e. and in measure (for each $k \geq 2$ and n such that $k(k-1)/2 < n \leq k(k+1)/2$,

$$\frac{k-1}{k+1} \chi_{[0,1]} \leq C_n f \leq \frac{k+1}{k-1} \chi_{[0,1]}.$$

The following example furnishes counterexamples to several likely-sounding conjectures.

EXAMPLE 2. The universal counterexample. Let D be the same as C except that $a_{11} = 0$. Let F be the family of all sequences of \mathfrak{X} -valued μ -measurable functions where $f \in F$ has the form $(f, -f, 0, 0, \dots)$, and $\mu(S) < \infty$. Then G , the scrambled closure of F , is the set of all sequences of the form $(f, g, 0, 0, \dots)$ and $P_1 F$ is the set of all sequences of the form $(f, 0, 0, \dots)$. Thus, each element of F , G , and $P_n F$ converges to 0 a.e. hence each element of DF , DG , and $DP_n F$ converges to 0 a.e. and since $\mu(S) < \infty$, the same is true of the other two modes of convergence. Df is identically 0 for every $f \in F$, so F is D -summable to 0 μ -uniformly in all three modes. It is not difficult to see that neither G nor $P_1 F$ is D -summable to 0 μ -uniformly in any of the three modes.

If $H = \{h\}$, $h = (h, -h, 0, 0, \dots)$, and $h(x) = x$ from the reals with Lebesgue measure to the reals, then H is closed under scrambling, Dh is identically 0, hence converges to 0 in measure, so H is D -summable to 0 μ -uniformly in measure, but $DP_1 h = (0, (1/2)h, (1/3)h, \dots)$ so does not converge to 0 in measure.

We wish to thank the referee for suggestions which led to the present form of Theorem I.3. Referring to Theorem I.3 for notation, our original theorem was that (a) and (d) \Rightarrow (f) which follows immediately from (C) and (T) of the present version. It was in investigating possible converse implications that the present form evolved.

The reader is invited to compare the strengths (as indicated in (T)) of the hypotheses and conclusions of implications in (B), (C), and (D) with those of the nonimplications in (B'), (C'), and (D') respectively. In particular, (C') shows that the rather unesthetic condition (d) is necessary for the implication (C) which is

the heart of this theorem; and the triviality of (B) makes the example of (B') all the more striking.

It follows from Theorem I.3 that if F is closed under scrambling and satisfies (a) (and (b) in case (ii)) then (d) \Leftrightarrow (f), answering a question raised by the referee.

I.3. THEOREM. *Let (S, Σ, μ) be a σ -finite measure space, \mathfrak{X} a normed linear space, T a lower triangular regular method of summability, F a family of sequences of μ -measurable \mathfrak{X} -valued functions, and G the scrambled closure of F . Fix one of the modes (i), (ii), or (iii) of convergence and let all convergence below be in that mode. Consider the following conditions:*

- (a) Each $g \in G$ is T -summable to 0.
- (b) For each n and each $f \in F$, $P_n f$ is T -summable to 0.
- (c) For each n and each $g \in G$, $P_n g$ is T -summable to 0.
- (d) For each n , $P_n F$ is T -summable to 0 μ -uniformly.
- (e) For each n , $P_n G$ is T -summable to 0 μ -uniformly.
- (f) F is T -summable to 0 μ -uniformly.
- (g) G is T -summable to 0 μ -uniformly.

Then the following implications hold:

- (T) (g) \Rightarrow (f), (g) \Rightarrow (a), (e) \Rightarrow (d), (e) \Rightarrow (c), (d) \Rightarrow (b), and (c) \Rightarrow (b).
- (A) (b) \Leftrightarrow (c) and (d) \Leftrightarrow (e).
- (B) (c) holds in cases (i) and (iii).
- (C) (a) and (d) \Rightarrow (g).
- (D) (g) \Rightarrow (e) in cases (i) and (iii), and (g) and (b) \Rightarrow (e) in case (ii).

Also, there are examples to demonstrate the following:

- (T') (f) $\not\Rightarrow$ (g), (a) $\not\Rightarrow$ (g), (c) $\not\Rightarrow$ (e), and (b) $\not\Rightarrow$ (d).
- (B') (g) $\not\Rightarrow$ (b) in case (ii).
- (C') (a) and (c) $\not\Rightarrow$ (f).
- (D') (a) and (c) and (f) $\not\Rightarrow$ (d) or (g).

Proof. The implications of (T) are trivial.

(A): The nontrivial directions in both of the implications in this part follow from the fact that for each fixed n and each $g \in G$, there exist $f^{(1)}, \dots, f^{(n)} \in F$ such that $P_n g = \sum_{i=1}^n P_i f^{(i)} - \sum_{i=2}^n P_{i-1} f^{(i)}$.

(B): For each n and each $g \in G$, $P_n g$ converges to 0 everywhere, therefore $TP_n g$ converges to 0 pointwise and therefore locally in measure.

(C): We give the details in case (i) and address occasional remarks to the other two cases which are similar.

For a sequence g of at least n functions, we use the notation $T_n g$ for $\sum_{j=1}^n a_{nj} g_j$ (where $T = (a_{ij})$). Note that for $f \in F$, $T_n f$ is the n th function in the sequence Tf .

Suppose that (C) is false. We will show that assuming (a), (d), and not (g) leads to a contradiction by constructing a sequence in G which is not T -summable to 0.

Since G is not T -summable to 0 μ -uniformly a.e., there exist $A \in \Sigma$ with $\mu(A) < \infty$.

and positive ε and α such that for every n , there exists $f \in G$ for which

$$\mu\left(\bigcup_{k=n}^{\infty} \{x \in A : \|(T_k f)(x)\| \geq 3\varepsilon\}\right) \geq 4\alpha.$$

(In cases (ii) and (iii), a similar statement is made for some $k \geq n$.)

Since the basis step in the following construction is just a simplified version of the induction step (with no previously constructed front end of the sequence to worry about), we give only the induction step.

Suppose we have chosen a finite sequence $g^{(n)}$ of functions g_1, \dots, g_n such that each $g_i = f_i$ for some $f \in F$, and there are integers $1 \leq S(1) < \dots < S(t) \leq n$ such that

$$\mu\left(\bigcup_{k=S(i)}^n \{x \in A : \|T_k g^{(n)}(x)\| \geq \varepsilon\}\right) \geq \alpha,$$

$i = 1, \dots, t$. The induction construction below will eventually yield a sequence g scrambled from F whose first n terms are $g^{(n)}$, so $T_k g = T_k g^{(n)}$ for $k \leq n$, but

$$\mu\left(\bigcup_{k=S(i)}^{\infty} \{x \in A : \|(T_k g)(x)\| \geq \varepsilon\}\right) \geq \mu\left(\bigcup_{k=S(i)}^n \{x \in A : \|(T_k g^{(n)})(x)\| \geq \varepsilon\}\right) \geq \alpha$$

for infinitely many n , and for infinitely many distinct $S(i)$ which will show that g does not converge to 0 μ -a.e. contradicting (a). The induction step proceeds as follows:

Let g' denote the sequence of functions whose first n members are those of $g^{(n)}$ and all the rest are 0. Then g' and hence Tg' converges to 0 a.e. (In cases (ii) and (iii) argue as in the proof of (A) to see that Tg' converges to 0 in mode (ii) or (iii).)

By (A), (d) \Rightarrow (e).

Choose $m > n$ so that

$$(1) \quad \mu\left(\bigcup_{k=m}^{\infty} \{x \in A : \|(T_k g')(x)\| \geq \varepsilon\}\right) \leq \alpha,$$

$$(2) \quad \mu\left(\bigcup_{k=m}^{\infty} \{x \in A : \|(T_k P_n f)(x)\| \geq \varepsilon\}\right) \leq \alpha$$

for all $f \in G$ (using the previous paragraph and (e)). Then choose $f \in G$ for which

$$\mu\left(\bigcup_{k=m}^{\infty} \{x \in A : \|(T_k f)(x)\| \geq 3\varepsilon\}\right) \geq 4\alpha.$$

Then there exists $l > m$ such that

$$\mu\left(\bigcup_{k=m}^l \{x \in A : \|(T_k f)(x)\| \geq 3\varepsilon\}\right) \geq 3\alpha.$$

(In cases (ii) and (iii) a similar statement is made for some $l \geq m$.)

If $g^{(l)}$ is that sequence of l functions, the first n of which agree with $g^{(n)}$ and the rest are taken from the proper place in the sequence f , and if $S(t+1) = m$, then the induction step is complete, since for $m \leq k \leq l$, $T_k g^{(l)} = T_k f + T_k g' - T_k P_n f$ so, except for two sets each of measure at most α (see (1) and (2) above), each point of

$$\bigcup_{k=m}^l \{x \in A : \|(T_k f)(x)\| \geq 3\varepsilon\}$$

is in

$$\bigcup_{k=m}^l \{x \in A : \|(T_k g^{(l)})(x)\| \geq \varepsilon\}$$

so the latter set has measure at least α , as was to be shown.

(D): (Sketch of proof.) By (A), (B), and (T), we may assume (c) in any case. Assuming (g), (c), and not (e) will lead to a contradiction to (g).

Pick the smallest n for which (e) fails. $T=(a_{ij})$, and for each j , $\lim_{i \rightarrow \infty} a_{ij}=0$. In any of cases (i), (ii), or (iii), it follows from the fact that $P_{n-1}G$ is T -summable to 0 μ -uniformly (if $n > 1$) but $P_n G$ is not that there exists $A \in \Sigma$ such that $\mu(A) < \infty$ in cases (i) or (iii) or $A=S$ in case (ii) and there exist ε and $\delta > 0$ and there exist infinitely many k and corresponding $g^{(k)} \in G$ such that

$$\mu\{x \in A : \|a_{kn} g_n^{(k)}(x)\| \geq \varepsilon\} \geq \delta.$$

Now fix $g \in G$. By using the facts g , $P_{n-1}g$, and $P_n g$ are T -summable to 0 and considering the subfamily of G formed by replacing g_n in g by $g_n^{(k)}$ for the infinitely many k described above, we see that this subfamily and hence G itself is not T -summable to 0 μ -uniformly, contradicting (g).

For the remainder of this proof, the symbols D , F , G , and H refer to the specific objects of I.2, Example 2.

(T'): The counterexamples called for here are all provided by this D , F , and G with $n=1$.

(B'): D and H satisfy (g), but (b) fails for $n=1$.

(C'): D and G satisfy (a) and (c), but not (f).

(D'): D and F satisfy (a), (c), and (f), but not (d) or (g).

The condition (d) of the previous theorem is cumbersome for applications. In what follows, we give some sufficient conditions for (d) to hold. These are notable in that they only depend on the family F and not the particular regular method of summability T .

I.4. THEOREM. *Let F be a family of sequences of \mathfrak{X} -valued μ -measurable functions. Suppose that for each n , and for each $A \in \Sigma$ of finite measure (in modes (I) and (III)) or for $A=S$ (in mode (II)), $\lim_{t \rightarrow \infty} \mu\{x \in A : \|f_n(x)\| \geq t\} = 0$ uniformly for $f \in F$. Then F satisfies hypothesis (d) of Theorem I.3 for any regular method of summability T .*

Proof. Fix $n \geq 1$, A as above, and $\varepsilon > 0$. In what follows,

$$\mu\{x \in A : \|f_k(x)\| \geq \varepsilon/0\} = 0.$$

As before, we detail the a.e. case. The other two are similar. Let $T=(a_{ij})$ and recall that for each j , $\lim_{i \rightarrow \infty} a_{ij}=0$.

Then for each $f \in F$,

$$\begin{aligned} \mu\left(\bigcup_{k=m}^{\infty} \{x \in A : \|(T_k P_n f)(x)\| \geq \varepsilon\}\right) &= \mu\left(\bigcup_{k=m}^{\infty} \left\{x \in A : \left\|\left(\sum_{j=1}^n a_{kj} f_j\right)(x)\right\| \geq \varepsilon\right\}\right) \\ &\leq \mu\left(\bigcup_{j=1}^n \bigcup_{k=m}^{\infty} \{x \in A : \|f_j(x)\| \geq \varepsilon/n |a_{kj}|\}\right) \\ &\leq \sum_{j=1}^n \mu\left\{x \in A : \|f_j(x)\| \geq \varepsilon/n \sup_{k \geq m} |a_{kj}|\right\} \end{aligned}$$

which converges to 0 as $m \rightarrow \infty$ uniformly for $f \in F$ by the “recall” above and the hypotheses of this theorem.

REMARK. The converse of the previous theorem holds in the sense that if the hypotheses fail for a family F , then for *some* regular method of summability, the conclusion fails. To see this, pick the smallest n for which the hypotheses fail, choose a suitable A involved in the failure for that n , and let T be any regular method of summability whose n th column contains infinitely many nonzero entries; (d) of Theorem I.3 will fail for F for this T and n .

I.5. THEOREM. Let F be a family of μ -measurable \mathfrak{X} -valued functions. Let $g: [0, \infty) \rightarrow [0, \infty)$ be increasing and suppose $\lim_{x \rightarrow \infty} g(x) = \infty$. Let $A \in \Sigma$. If

$$\{\text{ess sup } \{f(x) : x \in A\} : f \in F\} \quad \text{or} \quad \left\{ \int_A g(\|f(x)\|) d\mu(x) : f \in F \right\}$$

is a bounded set of real numbers, then $\lim_{t \rightarrow \infty} \mu\{x \in A : \|f(x)\| \geq t\} = 0$ uniformly for $f \in F$.

Proof. The case of uniformly bounded essential supremums is trivial. Suppose $\int_A g(\|f(x)\|) d\mu(x) \leq M$ for all $f \in F$. Pick $\varepsilon > 0$. Pick t so large that $t' \geq t$ implies that $M/g(t') < \varepsilon$. Since $\int_A g(\|f(x)\|) d\mu(x) \leq M$, $\mu\{x \in A : g(\|f(x)\|) \geq g(t')\} \leq M/g(t')$, for each $t' \geq t$, hence $\mu\{x \in A : \|f(x)\| \geq t'\} \leq \varepsilon$ for each $t' \geq t$. Since t depended only on ε (and g) and not f , the result follows.

For ease of future reference, we state the following:

I.6. COROLLARY. With notation as in the statement of Theorem I.3, if for each n , the collection $\{f_n : f \in F\}$ satisfies the hypotheses or conclusions of Theorem I.5 for all $A \in \Sigma$ of finite measure (modes (I) or (III)) or for $A = S$ (mode (II)), and if (a) holds, then (g) holds.

Proof. By Theorems I.4 or I.5, (d) holds, so by Theorem I.3 (C), (g) holds.

II. **Random variables and P -uniform convergence.** In this section we give the definitions of “ \mathfrak{X} -random variable” and related concepts and introduce uniform convergence concepts related to those of §I. Our principal theorem here, Theorem II.7, is the analog of Theorem I.3(C) for families of independent sequences of \mathfrak{X} -random variables.

II.1. DEFINITIONS. A probability space is a measure space (S, Σ, P) such that $P(S) = 1$. If \mathfrak{X} is a normed linear space, an \mathfrak{X} -random variable is a function $X: S \rightarrow \mathfrak{X}$ which is measurable (Definition I.1) where (S, Σ, P) is a probability space.

A finite collection X_1, \dots, X_n of \mathfrak{X} -random variables is independent if for each collection B_1, \dots, B_n of Borel sets in \mathfrak{X} ,

$$P\{x \in S : X_1(x) \in B_1 \ \& \ \dots \ \& \ X_n(x) \in B_n\} = \prod_{k=1}^n P\{x \in S : X_k(x) \in B_k\}.$$

A sequence $\{X_n\}$ of \mathfrak{X} -random variables is independent if each finite subsequence is.

The distribution of an \mathfrak{X} -random variable X is the measure $P_x(\cdot)$ on the Borel sets of \mathfrak{X} defined by

$$P_x(B) = P[X \in B] = P\{x \in S : X(x) \in B\}.$$

We will use the terminology from probability theory “almost surely” or “a.s.” for “almost everywhere” and “in probability” for “in measure”.

Since we wish to compare sequences of random variables with different domains, we extend our notion of uniform convergence slightly; this extension is possible since all measure spaces involved have total measure 1. If F is a family of sequences of random variables and if we denote the domain space of $X \in F$ by (S^X, Σ^X, P^X) , then the family F converges to 0

(I) P -uniformly almost surely if for each $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P^X \left(\bigcup_{k=n}^{\infty} \{x \in S^X : \|X_k(x)\| \geq \varepsilon\} \right) = 0$$

uniformly for $X \in F$, and

(II) P -uniformly in probability if for each $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P^X(\{x \in S^X : \|X_n(x)\| \geq \varepsilon\}) = 0$$

uniformly for $X \in F$.

If T is a regular method of summability and F is as above, then F is T -summable to 0 P -uniformly almost surely or in probability if the family $\{TX : X \in F\}$ converges to 0 P -uniformly almost surely or in probability respectively.

The following technical lemmas show that we are justified in treating the distributions of random variables rather than the variables themselves as the central concept here.

II.2. LEMMA. Let \mathfrak{X} and \mathfrak{X}' be normed linear spaces with Borel sets \mathcal{B} and \mathcal{B}' respectively. Let \mathfrak{X}^m be the cartesian product of m copies of \mathfrak{X} , and let \mathcal{B}^m be the σ -algebra of subsets of \mathfrak{X}^m generated by rectangles with Borel sides. Let $f: \mathfrak{X}^m \rightarrow \mathfrak{X}'$ be such that (a) $f^{-1}(B') \in \mathcal{B}^m$ for each $B' \in \mathcal{B}'$. Let X_1, \dots, X_m be an independent set of \mathfrak{X} -random variables. Then the distribution of $f(X_1, \dots, X_m)$ depends only on f and the distributions of X_1, \dots, X_m .

Proof. Exactly as in classical probability theory, if the joint distribution of X_1, \dots, X_m is defined on \mathcal{B}^m by

$$P_{X_1, \dots, X_m}(C) = P\{s \in S : (X_1(s), \dots, X_m(s)) \in C\},$$

then the distribution of $f(X_1, \dots, X_m)$ depends only on f and the joint distribution of X_1, \dots, X_m , and since X_1, \dots, X_m are independent,

$$P_{X_1, \dots, X_m} = P_{X_1} \times \cdots \times P_{X_m}.$$

II.3. LEMMA. *With the hypotheses of the previous lemma except that f is assumed continuous (with the product topology of \mathfrak{X}^m) instead of assuming (a), if \mathfrak{X} is separable, then (a) and therefore the conclusion of Lemma II.2 hold.*

Proof. Since \mathfrak{X} has a metric topology, \mathfrak{X} is second countable. In this case, it is well known and easily verified that \mathcal{B}^m is exactly the Borel sets of \mathfrak{X}^m , so (a) follows routinely.

II.4. COROLLARY. *Let X be an independent sequence of \mathfrak{X} -random variables and T a row finite matrix. Then for each n and ε the numbers*

$$P\left(\bigcup_{k=n}^{\infty} \{x \in S : \|(TX)_k(x)\| \geq \varepsilon\}\right) \quad \text{and} \quad P\{x \in S : \|(TX)_n(x)\| \geq \varepsilon\}$$

depend only on the distributions of X_1, X_2, \dots . In particular, whether a given family of independent sequences is T -summable to 0 P -uniformly almost surely or in probability depends only on the distributions of the random variables in the sequences.

Proof. Since each X_n is almost surely separably valued, its distribution is supported by a separable subspace of its range, so without loss of generality we assume that \mathfrak{X} is separable. Fix n, ε , and pick $m \geq n$. If $T = (a_{ij})$, there is an integer k such that $a_{ij} = 0$ for all i and j such that $1 \leq i \leq m$ and $j > k$. Define $f: \mathfrak{X}^k \rightarrow R$ (the real numbers) by

$$f(x_1, \dots, x_k) = \sup_{n \leq i \leq m} \|a_{i1}x_1 + \cdots + a_{ik}x_k\|.$$

Then f is continuous, so by Lemma II.3,

$$P\left(\bigcup_{i=n}^m \{x \in S : \|(TX)_i(x)\| \geq \varepsilon\}\right) = P[(f(X_1, \dots, X_k))^{-1}([\varepsilon, \infty))]$$

depends only on the distributions of the X_i 's. The corollary follows by taking $m = n$ for the second number and letting $m \rightarrow \infty$ and appealing to the countable additivity of P for the first number.

II.5. REMARKS. We now see that the hypothesis that \mathfrak{X} -random variables be essentially separably valued is necessary since otherwise linear combinations of \mathfrak{X} -random variables may not even be measurable (if the cardinality of \mathfrak{X} is greater than c , the diagonal of \mathfrak{X}^2 is not in \mathcal{B}^2). As we shall see in III.1, it is also necessary in order that an \mathfrak{X} -random variable have expectation.

In what follows we wish to speak of the class of "all \mathfrak{X} -random variables" or the class of "all independent sequences of \mathfrak{X} -random variables (which satisfy some restriction)". Now the class of domain spaces for "all \mathfrak{X} -random variables" would be "all probability spaces" and since every nonempty set can support at least one probability measure, included in consideration of "all \mathfrak{X} -random variables" is "all sets". This raises the specter of paradoxes of the type of Russell's paradox.

However, in the present situation, we are dealing with properties of \mathfrak{X} -random variables and independent sequences thereof which depend only on the distributions of the variables (a distribution on \mathfrak{X} is a measure μ on the Borel sets of \mathfrak{X} such that $\mu(\mathfrak{X})=1$ and μ has separable support; if μ is a distribution on \mathfrak{X} , the identity function from $(\mathfrak{X}, \mathcal{B}, \mu)$ to \mathfrak{X} is an \mathfrak{X} -random variable with distribution μ), and the set of all distributions on \mathfrak{X} , or the set of all sequences thereof, is not so large as to cause difficulties of the sort of Russell's paradox.

So while we will continue to speak in terms of random variables, more natural to the problems of pointwise convergence considered here, the remainder of this paper could be rewritten replacing "sets" S of sequences of random variables by sets S' of sequences of distributions and sentences of the form " $X \in S$ " by sentences of the form "the sequence of distributions of variables of X belongs to S' " with only technical changes.

II.6. DEFINITIONS. Let F be a family of independent sequences of \mathfrak{X} -random variables. An independent sequence Y of \mathfrak{X} -random variables is an *independent sequence scrambled from F* if for each n there exists $X \in F$ such that X_n and Y_n have the same distribution. F is *closed under independent scrambling* if each independent sequence scrambled from F is in F . The *independent scrambled closure of F* is the set of all independent sequences scrambled from F . Clearly, F is closed under independent scrambling if and only if F is the independent scrambled closure of F .

II.7. THEOREM. Let \mathfrak{X} be a normed linear space, T a lower triangular regular method of summability, F a family of independent sequences of \mathfrak{X} -random variables, and G the independent scrambled closure of F . Fix one of the modes "almost surely" or "in probability" and let all convergence below be in that mode. Then if each element of G is T -summable to 0 and for each n the family $P_n F$ is T -summable to 0 P -uniformly, then F and G are T -summable to 0 P -uniformly.

Proof. For each $X \in F$ let (S^X, Σ^X, P^X) be its domain and let (S', Σ', P') be the product space $\prod_{X \in F} (S^X, \Sigma^X, P^X)$. For each $X \in F$, let X' be the sequence of \mathfrak{X} -random variables defined on (S', Σ', P') by $X'(s') = X(s'(X))$. Let $F' = \{X' : X \in F\}$ and let G' be the scrambled closure of F' .

Each X in F has the same sequence of distributions as the corresponding X' in F' , so by Corollary II.4 and the assumption that $P_n F$ is T -summable to 0 P -uniformly, it follows that $P_n F'$ is T -summable to 0 P -uniformly. Also, each Y in

G' is an independent scrambled sequence from F , therefore $Y \in G$ so Y is T -summable to 0. We have verified (a) and (d) of Theorem I.3 for the families F' and G' so by Theorem I.3(C), F' and G' are T -summable to 0 P -uniformly, hence again by Corollary II.4, F and G are T -summable to 0 P -uniformly.

II.8. REMARKS. Theorem III.2 of [4] can be improved to the following: if for fixed $k \geq 2$ and $\varepsilon > 0$, F is any family of independent sequences of random variables such that the expectation of each random variable is 0 and the variance of each random variable is at most 1 and each sequence takes its values in a k, ε -convex space, then the family F is $(C, 1)$ -summable to 0 P -uniformly almost surely. This follows from Theorem II.7, and the argument is typical of applications we have for this theorem. For each $X \in F$, let \mathfrak{X}^X be the space in which random variables in X take their values and let \mathfrak{X}' be the Hilbert sum of the \mathfrak{X}^X . By Theorem II.17 and Example I.3(ii) of [4], \mathfrak{X}' is B -convex, and \mathfrak{X}' contains copies of each \mathfrak{X}^X in a natural way. Let ι_X be the natural injection of \mathfrak{X}^X into \mathfrak{X}' and for each $X \in F$, let $X' = \iota_X X$. By [1, Theorem 6] each sequence from $F' = \{X' : X \in F\}$ and each independent scrambled sequence from F' is $(C, 1)$ -summable to 0 a.s. and Theorem I.5 applies with $g(x) = x^2$ to show that Theorem II.7 applies to the family F' and hence F is $(C, 1)$ -summable to 0 P -uniformly almost surely.

III. **Vector-valued strong laws of large numbers.** We deal here with random variables which take their values in a normed linear space and of necessity we must make sense of the notion of "expectation" for them. This involves the integration of functions from a measure space to a normed linear space. A general reference for this topic is Dunford and Schwartz [2], Chapter III. We outline the necessary notions below.

III.1. *Definitions and preliminary remarks.* Let \mathfrak{X} be a normed linear space and let X be an \mathfrak{X} -random variable (Definition II.1). Then X is measurable in terms of Dunford and Schwartz [2], Chapter III (by Theorem III.6.10 there). By [2], III.2.22(a), the expectation (integral) of an \mathfrak{X} -random variable X exists if and only if the integral $\int_{\Omega} \|X(\omega)\| dP(\omega)$ is finite, in which case the expectation of X is an element y of the completion \mathfrak{X}^c of \mathfrak{X} characterized by the fact that $x^*(y) = \int_{\Omega} x^* X dP$ for all $x^* \in \mathfrak{X}^*$, the conjugate space of \mathfrak{X} ([2], III.2.19(c); recall that \mathfrak{X}^* separates points of \mathfrak{X}^c so that at most one y satisfies these restrictions). The expectation of X is denoted by $E(X)$. For an \mathfrak{X} -random variable X with expectation, the variance is $\sigma^2(X) = \int_{\Omega} \|X - E(X)\|^2 dP$, and the standard deviation is the nonnegative square root of variance, $\sigma(X)$. If $x \in \mathfrak{X}$, and X is an \mathfrak{X} -random variable, then X and $X - x$ have the same variance and standard deviation, so in the following, it will be no loss of generality to consider \mathfrak{X} -random variables centered at expectation (with expectation 0). Denote the essential supremum of X by $\beta(X)$. If $E(X) = 0$, $\sigma(X) \leq \beta(X)$.

An \mathfrak{X} -random variable X is symmetric if for each Borel set B of X ,

$$PX^{-1}(B) = PX^{-1}(-B), \quad \text{where } -B = \{-x : x \in B\}.$$

If $\{z_n\}$ is a sequence in a linear space, then $C\{z_n\}$ is the sequence of Cesaro $(C, 1)$ means:

$$C\{z_n\} = \{z_1, (z_1 + z_2)/2, (z_1 + z_2 + z_3)/3, \dots\}.$$

Note that C is a lower triangular regular method of summability (Definition I.1), a fact which will be needed below. If P is a property of sequences, then we say that a sequence z is $(C, 1)P$ if the sequence Cz has property P . A real sequence $\{a_n\}$ is a K -sequence (K for Kolmogorov) if $\sum_{n=1}^{\infty} a_n^2/n^2 < \infty$. Let \mathbf{K} be the set of all nonnegative valued K -sequences, \mathbf{CB} the set of all nonnegative valued $(C, 1)$ bounded sequences, \mathbf{CN} the set of all nonnegative valued $(C, 1)$ null sequences, and \mathbf{J} the set of all nonnegative valued sequences \mathbf{a} such that $(a_1^2 + \dots + a_n^2)/n^2 \rightarrow 0$ as $n \rightarrow \infty$.

Just as in classical probability theory, if X is an \mathfrak{X} -random variable, $E(X)$, $\sigma^2(X)$, $\sigma(X)$, $\beta(X)$ and whether or not X is symmetric are all determined by the distribution of X , as is the essential cardinality of the range of X (if the domain of X is (S, Σ, P) , the essential cardinality of the range of X is the infimum of the cardinalities of all sets of the form $X(S \setminus N)$ where $N \in \Sigma$ with $P(N) = 0$).

Let $\mathcal{S}(\mathfrak{X})$ be the collection of all independent sequences of \mathfrak{X} -random variables such that each variable in a sequence of $\mathcal{S}(\mathfrak{X})$ has expectation 0. $\mathcal{S}_s(\mathfrak{X})$ is the subset of $\mathcal{S}(\mathfrak{X})$ in which each random variable is symmetric. $\mathcal{S}_2(\mathfrak{X})$ and $\mathcal{S}_c(\mathfrak{X})$ are the subsets of $\mathcal{S}(\mathfrak{X})$ in which each random variable has an essential range of respectively at most 2 and countably many elements. Let \mathbf{a} be a nonnegative real sequence. $\mathcal{S}_\beta(\mathfrak{X}, \mathbf{a})$ and $\mathcal{S}_\sigma(\mathfrak{X}, \mathbf{a})$ are the sets of all sequences $\{X_n\}$ in $\mathcal{S}(\mathfrak{X})$ for which, respectively, $\beta(X_n) \leq a_n$ and $\sigma(X_n) \leq a_n$. The symbol " \mathcal{S} " with more than one of the above subscripts and with appropriate arguments stands for the intersection of the sets involving the subscripts used. " \mathfrak{X} " and " \mathbf{a} " will be dropped where no confusion can result. The inclusions

$$\mathcal{S}_2(\mathfrak{X}) \subset \mathcal{S}_{cs}(\mathfrak{X}) \subset \mathcal{S}_s(\mathfrak{X}) \subset \mathcal{S}(\mathfrak{X})$$

and $\mathcal{S}_\beta(\mathfrak{X}, \mathbf{a}) \subset \mathcal{S}_\sigma(\mathfrak{X}, \mathbf{a})$ are formal. Since $\mathcal{S}_{2\sigma}(\mathfrak{X}, \mathbf{a}) = \mathcal{S}_{2s\beta}(\mathfrak{X}, \mathbf{a}) =$ all sets formally lying between these two, we use the notation $\mathcal{S}_2(\mathfrak{X}, \mathbf{a})$ to denote this set.

A sequence from $\mathcal{S}(\mathfrak{X})$ satisfies the strong law of large numbers (SLLN) if it is almost surely $(C, 1)$ null (in the norm topology of \mathfrak{X}). If L is a list of subscripts from the set $\{2, c, s, \beta, \sigma\}$ including at least one of 2, β , and σ , then $\mathcal{A}_L(\mathfrak{X})$ is the set of all nonnegative sequences \mathbf{a} such that every element of $\mathcal{S}_L(\mathfrak{X}, \mathbf{a})$ satisfies SLLN.

A family $\mathcal{F} \subset \mathcal{S}(\mathfrak{X})$ satisfies SLLN P -uniformly if it is C -summable to 0 P -uniformly almost surely (Definition II.1).

For L and \mathbf{a} as above and for $X \in \mathcal{S}(\mathfrak{X})$, whether or not X is in $\mathcal{S}_L(\mathfrak{X}, \mathbf{a})$ is completely determined by the distributions of the variables in X , so membership in $\mathcal{S}_L(\mathfrak{X}, \mathbf{a})$ is a restriction of the type mentioned in the first proposition stated in the introduction to this paper. By Corollary II.4, whether or not X satisfies SLLN depends only on the distributions of the variables in X . Thus, identification of $\mathcal{A}_L(\mathfrak{X})$ or any portion thereof is a strong law of large numbers for \mathfrak{X} -random variables.

In the set of all real sequences, we define addition, scalar multiplication, and order componentwise. The least upper bound of two elements x and y is denoted by $x \vee y$. Let A be a set of nonnegative valued sequences. A is closed under sums if $x + y \in A$ for all x and y in A . A is closed under nonnegative multiples if $\alpha x \in A$ for all $\alpha \geq 0$ and all x in A . A is closed under maxima if $x \vee y \in A$ for all x and y in A . A is closed under subdominance if x in A and $0 \leq z_n \leq x_n$ for all n implies that z is in A . A is a proper cone if it is closed under all of these. Since $x \vee y \leq x + y \leq 2(x \vee y)$ for all nonnegative sequences x and y , in the presence of the other two properties, A is closed under sums if and only if it is closed under maxima.

III.2. THEOREM. *J, K, CB, and CN are proper cones. If \mathfrak{X} is a normed linear space, the sets A_L of Definition III.1 are proper cones.*

Proof. We use the fact that if $\{X_n\} \in S_L(a + b)$, then there exist $\{Y_n\} \in S_L(a)$ and $\{Z_n\} \in S_L(b)$ such that for each n , $X_n = Y_n + Z_n$ ($Y_n = a_n X_n / (a_n + b_n)$ if the denominator is nonzero, etc.). The rest follows routinely.

III.3. THEOREM. *Let $\mathfrak{Y} \subset \mathfrak{X}$ and L be as in Definition III.1. Then $A_L(\mathfrak{X}) \subset A_L(\mathfrak{Y})$. $A_L(\mathfrak{X})$ is invariant under equivalent renorming of \mathfrak{X} .*

Proof. For each a , $S_L(\mathfrak{Y}, a) \subset S_L(\mathfrak{X}, a)$, so the inclusion follows. Renorming \mathfrak{X} with an equivalent norm does not change the topology, so $S(\mathfrak{X})$ is the same, as are those elements of $S(\mathfrak{X})$ which satisfy SLLN. While $\beta(X)$ and $\sigma(X)$ may change, this change is by a uniformly bounded factor, so by Theorem 2, $A_L(\mathfrak{X})$ is not changed.

III.4. THEOREM. *Let \mathfrak{X} be a normed linear space. Then*

$$A_\sigma \subset A_{\sigma s} \subset A_{\beta s}, \quad A_\sigma \subset A_\beta \subset A_{\beta s}, \quad A_{\beta s} \subset A_{c\beta s} \subset A_2.$$

Proof. Each inclusion follows since the criterion for membership in each alleged larger set is less restrictive in that a smaller set from $S(\mathfrak{X})$ is required to satisfy SLLN.

The next several theorems are devoted to the question of which of the above inclusions can be reversed, and when they cannot, how much along that line can be said. The following technical result is needed.

III.5. THEOREM. *Let \mathfrak{X} be a normed linear space, let L be as in Definition III.1 and let $a \in A_L(\mathfrak{X})$. Then $S_L(\mathfrak{X}, a)$ satisfies SLLN P -uniformly.*

Proof. By definition of $A_L(\mathfrak{X})$, every sequence of $S_L(\mathfrak{X}, a)$ is C -summable to 0 almost surely, and the form of $S_L(\mathfrak{X}, a)$ insures that it is closed under independent scrambling. Since for each fixed n , the n th terms of the sequences of $S_L(\mathfrak{X}, a)$ are uniformly bounded in essential supremum by a_n or in variance by a_n^2 , Theorem I.5 applies and so by Theorem II.7, the present result follows.

III.6. THEOREM. *Let \mathfrak{X} be a normed linear space. Then $A_2(\mathfrak{X}) = A_{c\beta s}(\mathfrak{X})$.*

Proof. One inclusion is established by Theorem III.4. Let \mathbf{a} be an arbitrary element of A_2 . By Theorem III.5, $S_2(\mathbf{a})$ satisfies SLLN P -uniformly, so there exists for each $n \geq 1$ and $\varepsilon > 0$ a nonnegative number $\eta(n, \varepsilon)$ with the property that for each $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \eta(n, \varepsilon) = 0$ and such that if $\{X_n\} \in S_2(\mathbf{a})$, then

$$P\left(\bigcup_{k=n}^{\infty} \{\omega \in \Omega : \|X_1(\omega) + \cdots + X_k(\omega)\|/k \geq \varepsilon\}\right) \leq \eta(n, \varepsilon).$$

Then, for each such $\{X_n\}$ and each $m \geq n$,

$$P\left(\bigcup_{k=n}^m \{\omega \in \Omega : \|X_1(\omega) + \cdots + X_k(\omega)\|/k \geq \varepsilon\}\right) \leq \eta(n, \varepsilon).$$

We shall verify that each sequence of $S_{c\beta s}(\mathbf{a})$ also satisfies this last inequality for every $m \geq n$ so that it also satisfies the second last inequality, therefore it satisfies SLLN, and it will follow that $\mathbf{a} \in A_{c\beta s}$.

Now let $\{Y_n\} \in S_{c\beta s}(\mathbf{a})$. Since each Y_n is symmetric, if x is in the range of Y_n , so is $-x$, and $P[Y_n = x] = P[Y_n = -x]$ and we may assume without loss of generality that these probabilities are nonzero. Denote the domain of the Y_n 's by (Ω, Σ, P) . Let $Y_n = \pm y_{n1}$ on A_{n1} , $= \pm y_{n2}$ on A_{n2} , and so forth, where for $i \neq j$, $y_{ni} \neq \pm y_{nj}$ and (by perhaps discarding a null subset of Ω) the sets A_{n1}, A_{n2}, \dots form a measurable partition of Ω into sets of nonzero measure. $\|y_{ni}\| \leq a_n$ for each n and i . Some y_{ni} may be 0. Now fix $n, m \geq n$, and $\varepsilon > 0$. For each j between 1 and m , let A_j be one of the sets A_{ji} (possibly different i for different j) and let $A = \bigcap_{j=1}^m A_j$. Since the sets A_1, \dots, A_m are independent and of positive probability, $P(A) > 0$. The sets A formed in this way by all possible choices of the A_j 's form a countable measurable partition of Ω . Let (A, Σ_A, P_A) be the probability space defined by $\Sigma_A = \{B \in \Sigma : B \subset A\}$ and for $B \in \Sigma_A$, $P_A(B) = P(B)/P(A)$. Using the fact that each A_j is the inverse image under Y_j of a Borel set, it follows that the sequence $\{Y_j|A : 1 \leq j \leq m\}$ of Y_j 's restricted to this probability space is independent, so is the initial segment (fill out from $m+1$ onward with 0's) of a sequence in $S_2(\mathbf{a})$. Therefore,

$$P_A\left(\bigcup_{k=n}^m \{\omega \in A : \|Y_1(\omega) + \cdots + Y_k(\omega)\|/k \geq \varepsilon\}\right) \leq \eta(n, \varepsilon),$$

and therefore

$$P\left(\bigcup_{k=n}^m \{\omega \in A : \|Y_1(\omega) + \cdots + Y_k(\omega)\|/k \geq \varepsilon\}\right) \leq P(A)\eta(n, \varepsilon).$$

Summing this last inequality over all A 's described above gives us the desired result.

III.7. THEOREM. *Let \mathfrak{X} be a normed linear space. Then $A_{c\beta s}(\mathfrak{X}) = A_{\beta s}(\mathfrak{X})$.*

This result depends on the following approximation:

III.8. LEMMA. *Let X be a symmetric \mathfrak{X} -random variable and let $\varepsilon > 0$. Then there exists a symmetric \mathfrak{X} -random variable Y with countable range such that $\|X - Y\| \leq \varepsilon$*

almost surely, and for each Borel set B of \mathfrak{X} there is a Borel set B' of \mathfrak{X} such that $Y^{-1}(B) = X^{-1}(B')$.

Proof. Since X is almost surely separably valued, there exists a countable sequence $\{x_n\}$ which is almost surely dense in the range of X , with $x_0 = 0$. For $n \geq 0$, let B_n be the ball in \mathfrak{X} with center x_n and radius ε . Let $K_0 = B_0$,

$$K_n^+ = B_n \setminus \left[K_0 \cup \bigcup_{k=1}^{n-1} (K_k^+ \cup K_k^-) \right],$$

and $K_n^- = -K_n^+$. It is routine to verify that the sequence K_0, K_1^+, K_1^-, \dots is pairwise disjoint. Let $Y(\omega) = 0$ if $X(\omega) \in K_0$, $Y(\omega) = x_n$ if $X(\omega) \in K_n^+$, and $Y(\omega) = -x_n$ if $X(\omega) \in K_n^-$, and let $Y(\omega) = 0$ otherwise, this last occurring on a set of probability 0. Then symmetry of X implies that of Y , and an inspection of the construction shows that the other claims follow.

Proof of Theorem 7. One inclusion is contained in Theorem III.4. Now let $\mathbf{a} \in A_{\varepsilon\beta\sigma}$ and $\{X_n\} \in S_{\beta\sigma}(\mathbf{a})$. We will be done when we show that $\{X_n\}$ satisfies SLLN. Fix $\varepsilon > 0$.

If $a_n = 0$, then $X_n = 0$ a.s. so let $Y_n = 0$. Otherwise, apply Lemma III.8 to X_n with $\varepsilon_n = \min(\varepsilon, a_n)$. By the last statement of Lemma III.8, the sequence $\{Y_n\}$ is independent, and since $\|X_n - Y_n\| \leq a_n$ a.s. and $\beta(X_n) \leq a_n$, $\beta(Y_n) \leq 2a_n$, so by the choice of \mathbf{a} and Theorem III.2, $\{Y_n\}$ satisfies SLLN. This is equivalent to saying

$$\text{ess sup}_{\Omega} \limsup_{n \rightarrow \infty} n^{-1} \left\| \sum_{k=1}^n Y_k \right\| = 0.$$

Abbreviating “ $\text{ess sup}_{\Omega} \limsup_{n \rightarrow \infty}$ ” by “ L ”, we then have

$$Ln^{-1} \left\| \sum_{k=1}^n X_k \right\| \leq Ln^{-1} \left\| \sum_{k=1}^n Y_k \right\| + Ln^{-1} \sum_{k=1}^n \|X_k - Y_k\| < \varepsilon.$$

Since ε was arbitrary, it follows that $Ln^{-1} \left\| \sum_{k=1}^n X_k \right\| = 0$ and we are done.

III.9. THEOREM. *Let \mathfrak{X} be a normed linear space. Then*

$$A_{\beta}(\mathfrak{X}) = A_{\beta\sigma}(\mathfrak{X}) \quad \text{and} \quad A_{\sigma}(\mathfrak{X}) = A_{\sigma\sigma}(\mathfrak{X}).$$

Proof. As before, Theorem III.4 gives one direction of each equality claimed. We follow the technique of [1], Theorem 10. The proofs for the two cases are nearly identical in form. We give the β case and indicate any necessary changes for the σ case as we go along.

If $\mathfrak{X} = \{0\}$, then $A_L(\mathfrak{X}) =$ all nonnegative valued real sequences for each L of Definition 1 and the theorem follows. In what follows, suppose that \mathfrak{X} is nontrivial.

Let $\mathbf{a} \in A_{\beta\sigma}$ and let $\{X_n\} \in S_{\beta}(\mathbf{a})$. As before, if we can show that $\{X_n\}$ satisfies SLLN, we are done. Let (Ω, Σ, P) be the domain of the X_n . Define Y_n on the product space $\Omega \times \Omega$ by $Y_n(\omega_1, \omega_2) = X_n(\omega_1) - X_n(\omega_2)$. Then $\{Y_n\} \in S_{\sigma}$ and easily, $\beta(Y_n) \leq 2\beta(X_n)$. (By the brute force approximation of [1, p. 41], it follows that $\sigma(Y_n) \leq 2\sigma(X_n)$.) It follows by the choice of \mathbf{a} and Theorem III.2 that $\{Y_n\}$ satisfies SLLN, so that on a subset of $\Omega \times \Omega$ of probability 1, the $(C, 1)$ means of $\{Y_n\}$

converge to 0. Almost every section of that subset has probability 1, so there exist $\omega_0 \in \Omega$ and $S \in \Sigma$ such that $P(S) = 1$ and for all $\omega \in S$,

$$\lim_{n \rightarrow \infty} n^{-1} \left\| \sum_{k=1}^n X_k(\omega) - X_k(\omega_0) \right\| = 0.$$

So if we can show that the sequence $X_n(\omega_0)$ is $(C, 1)$ null, we are done.

Fix $\varepsilon > 0$. Let $S_n = \{\omega \in S : n^{-1} \|\sum_{k=1}^n X_k(\omega) - X_k(\omega_0)\| < \varepsilon\}$. Then $\lim_{n \rightarrow \infty} P(S_n) = 1$, so there exists N_1 such that $n \geq N_1$ implies $P(S_n) > 1/2$.

Now consider \mathfrak{X} as a real normed linear space and let x^* be a real-valued continuous linear functional on \mathfrak{X} such that $\|x^*\| \leq 1$. The sequence $\{x^* X_n\} \in \mathcal{S}_\beta(R, \mathbf{a})$ (R is the real numbers). We now need the following.

III.10. LEMMA. $K \subset J$. If \mathfrak{X} is a nontrivial normed linear space, $A_{\beta s}(\mathfrak{X}) \subset J$ and $A_{\sigma s}(\mathfrak{X}) \subset K$.

Proof. The first claim follows immediately from the Kronecker Lemma (see [5, p. 238]). Since R is essentially a subspace of \mathfrak{X} , it is sufficient by Theorem III.3 to prove this lemma for $\mathfrak{X} = R$. But the third inclusion claimed is well known in this case (Feller in [3] relegates it to an exercise; X.8.19, p. 247). For the second inclusion claimed, first note that $\mathbf{a} \in A_{\beta s}(R)$ implies that $a_n < n$ with at most finitely many exceptions, for if $\{Z_n\} \in \mathcal{S}_{\beta s}(R, \mathbf{a})$ is defined by the requirement that $Z_n = \pm a_n$ with equal probability, then $P\{n^{-1}|Z_1 + \dots + Z_n| \geq 1\} \geq 1/2$, for each n for which $a_n \geq n$. Since this same sequence $\{Z_n\}$ must satisfy the weak law of large numbers, it must satisfy the necessary conditions of the Classical Degenerate Convergence Criterion ([5, p. 278]), and condition (iii) there, with the symmetry of the Z_n , gives us the desired result.

Proof of Theorem 9 (continued). Note that

$$\sigma^2\left(n^{-1} \sum_{k=1}^n x^* X_k\right) = n^{-2} \sum_{k=1}^n \sigma^2(x^* X_k) \leq n^{-2} \sum_{k=1}^n \sigma^2(X_k) \leq (a_1^2 + \dots + a_n^2)/n^2$$

in either case β or case σ . By Lemma III.10, there exists N_2 such that if $n \geq N_2$, $(a_1^2 + \dots + a_n^2)/\varepsilon^2 n^2 < 1/2$. Note that while N_2 depends on \mathbf{a} and ε , it does not depend on x^* .

Let $N = \max(N_1, N_2)$ and fix $n \geq N$. Chebyshev's inequality tells us that

$$P\left(n^{-1} \left| \sum_{k=1}^n x^* X_k \right| \geq \varepsilon\right) < 1/2.$$

Hence there is a point ω_1 in S_n for which $n^{-1}|x^* X_1(\omega_1) + \dots + x^* X_n(\omega_1)| < \varepsilon$. It then follows that

$$\begin{aligned} n^{-1} \left| \sum_{k=1}^n x^*(X_k(\omega_0)) \right| &\leq n^{-1} \left| \sum_{k=1}^n x^*(X_k(\omega_1) - X_k(\omega_0)) \right| + n^{-1} \left| \sum_{k=1}^n x^*(X_k(\omega_1)) \right| \\ &\leq n^{-1} \left\| \sum_{k=1}^n (X_k(\omega_1) - X_k(\omega_0)) \right\| + \varepsilon \\ &\leq 2\varepsilon. \end{aligned}$$

The inequality between the extreme ends of the above chain holds simultaneously for all x^* in the unit ball of \mathfrak{X}^* . It follows by [2], Corollary II.3.15 (a corollary to the Hahn-Banach theorem) that $n^{-1}\|X_1(\omega_0) + \dots + X_n(\omega_0)\| \leq 2\epsilon$, and this holds for all $n \geq N$. This was what we needed to complete this proof.

This leaves only one inclusion from Theorem III.4 which we have not reversed. While we cannot reverse it (Theorem III.13) we have the following.

III.11. THEOREM. *Let \mathfrak{X} be a nontrivial normed linear space. Then*

$$A_{\beta s}(\mathfrak{X}) \cap K \cap CB \subset A_{\sigma s}(\mathfrak{X}) \subset A_{\beta s}(\mathfrak{X}) \cap K.$$

Proof. The second inclusion claimed follows from Theorem III.4 and Lemma III.10. For the first inclusion, we proceed as usual. Let \mathbf{a} be a $(C, 1)$ bounded K -sequence in $A_{\beta s}$. Let $\{X_n\} \in S_{\sigma s}(\mathbf{a})$. We must show that $\{X_n\}$ satisfies SLLN. Let “ L ” be the operator defined in the proof of Theorem III.7. The argument that follows is essentially that of [1, Lemma 9].

Fix $M > 0$. Define for $n \geq 1$, $Y_n = X_n$, $Z_n = 0$ wherever $X_n \leq Ma_n$, $Y_n = 0$, $Z_n = X_n$ wherever $X_n > Ma_n$. Then $E(Ma_n\|Z_n\|) \leq E(\|Z_n\|^2) \leq E(\|X_n\|^2) = \sigma^2(X_n) \leq a_n^2$, for each n , so $E(\|Z_n\|) \leq a_n/M$. Let $\alpha = \limsup C(\mathbf{a})$, finite by assumption. $\{Y_n\} \in S_{\beta s}(M\mathbf{a})$ and $\{\|Z_n\|\}$ is an independent sequence. Furthermore, since $\{\|Z_n\|\}$ satisfies the inequalities above and \mathbf{a} is a K -sequence, the strong law of large numbers for real random variables implies that

$$\begin{aligned} L\|C\{Z_n\}\| &\leq LC\{\|Z_n\|\} \\ &\leq LC\{(\|Z_n\| - E(\|Z_n\|))\} + \limsup_n C\{E(\|Z_n\|)\} \\ &\leq 0 + \limsup_n C(\mathbf{a})/M \\ &\leq \alpha/M. \end{aligned}$$

The choice of \mathbf{a} and Theorem III.2 guarantee that $L\|C\{Y_n\}\| = 0$. Thus $L\|C\{X_n\}\| \leq L\|C\{Y_n\}\| + L\|C\{Z_n\}\| \leq \alpha/M$. Since this is true for all $M > 0$, it follows that $\{X_n\}$ satisfies SLLN and we are done.

III.12. REMARKS. We now have $A_{\sigma s}(\mathfrak{X}) = A_{\sigma}(\mathfrak{X})$, $A_2(\mathfrak{X}) = A_{c\beta s}(\mathfrak{X}) = A_{\beta s}(\mathfrak{X}) = A_{\beta}(\mathfrak{X})$, and $A_{\beta}(\mathfrak{X}) \cap CB \cap K \subset A_{\sigma}(\mathfrak{X}) \subset A_{\beta}(\mathfrak{X}) \cap K$. It will follow from Theorem III.13 that the inclusion $A_{\sigma}(\mathfrak{X}) \subset A_{\beta}(\mathfrak{X})$ is always proper for nontrivial \mathfrak{X} since $A_{\sigma}(\mathfrak{X}) \subset K$, $CN \subset A_{\beta}(\mathfrak{X})$ and there exists a $(C, 1)$ null sequence which is not a K -sequence.

Theorem III.13 is the theorem implicitly promised by the second part of the title of this paper. Its assertions include the following:

If $\{X_n\}$ is an independent sequence of \mathfrak{X} -random variables such that $E(X_n) = 0$ for each n and $n^{-1}(\text{ess sup } \|X_1\| + \dots + \text{ess sup } \|X_n\|) \rightarrow 0$ as $n \rightarrow \infty$, then $\{X_n\}$ satisfies SLLN.

If $\{X_n\}$ is an independent sequence of \mathfrak{X} -random variables such that $E(X_n) = 0$ for each n , $n^{-1}(\sigma(X_1) + \dots + \sigma(X_n)) \rightarrow 0$ as $n \rightarrow \infty$, and $\sum_{n=1}^{\infty} \sigma^2(X_n)/n^2 < \infty$, then $\{X_n\}$ satisfies SLLN.

If the restriction on $\{\text{ess sup } \|X_n\|\}$ in the first statement is weakened, or if the restrictions on $\{\sigma(X_n)\}$ in the second statement are weakened, the resulting statement is no longer true for all normed linear spaces.

III.13. THEOREM. *Let \mathfrak{X} be a normed linear space. Then $CN \subset A_\beta(\mathfrak{X})$ and $CN \cap K \subset A_\sigma(\mathfrak{X})$. For l_1 , these inclusions are equalities. The inclusion $CN \cap K \subset CN$ is proper.*

Proof. Since $CN \subset CB$, the first inclusion and Theorem III.11 establish the second inclusion. If $\mathbf{a} \in CN$ and $\{X_n\} \in \mathcal{S}_\beta(\mathbf{a})$, then by the triangle inequality in \mathfrak{X} , $C\{X_n\}$ converges to 0 almost surely, so $\mathbf{a} \in A_\beta(\mathfrak{X})$. For l_1 , the first equality with Theorem III.11 and Lemma III.10 establishes the second equality. For the first equality, let $\mathbf{a} \notin CN$ and let $\{x_n\}$ denote the usual basis of l_1 . Define $\{X_n\} \in \mathcal{S}_2(l_1)$ by requiring that $X_n(\Omega) = \{\pm a_n x_n\}$. Then $\|X_1 + \cdots + X_n\|/n = (a_1 + \cdots + a_n)/n$ everywhere, so $\{X_n\}$ does not satisfy SLLN.

To verify the last claim, we construct a $(C, 1)$ null sequence \mathbf{a} which is not a K -sequence. This sequence has many terms equal to 0. The k th nonzero term of \mathbf{a} occurs at position $n(k)$, and $a_{n(k)} = n(k)/k^{1/2}$. This contributes $1/k$ to $\sum_{n=1}^{\infty} a_n^2/n^2$ while the increase of $C(\mathbf{a})$ from $n(k) - 1$ to $n(k)$ is at most $1/k^{1/2}$. Independent of the choice of the $n(k)$'s, this insures that the resulting sequence is not a K -sequence. To insure that \mathbf{a} is a $(C, 1)$ null sequence, choose $n(k)$ so large that $C(\mathbf{a})$ at $n(k) - 1$ is at most $1/k^{1/2}$.

REFERENCES

1. Anatole Beck, *On the strong law of large numbers. Ergodic theory*, Proc. Internat. Sympos. (Tulane Univ., New Orleans, La., 1961), Academic Press, New York, 1963, pp. 21–53. MR 28 #3470.
2. Nelson Dunford and Jacob T. Schwartz, *Linear operators. I, General theory*, Pure and Appl. Math., vol. 7, Interscience, New York, 1958. MR 22 #8302.
3. William Feller, *An introduction to probability theory and its applications*. Vol. 1, Wiley, New York, 1950. MR 12, 424.
4. Daniel P. Giesy, *On a convexity condition in normed linear spaces*, Trans. Amer. Math. Soc. **125** (1966), 114–146. MR 34 #4866.
5. Michel Loève, *Probability theory*, Van Nostrand, Princeton, N. J., 1963. MR 34 #3596.
6. E. Parzen, *Some conditions for uniform convergence of integrals*, Proc. Amer. Math. Soc. **5** (1954), 55–58. MR 15, 514.
7. ———, *On uniform convergence of families of sequences of random variables*, Univ. California Publ. Statist. **2** (1954), 23–53. MR 15, 806.

UNIVERSITY OF WISCONSIN,
MADISON, WISCONSIN
UNIVERSITY OF SOUTHERN CALIFORNIA,
LOS ANGELES, CALIFORNIA