A CONVEXITY CONDITION IN BANACH SPACES AND
THE STRONG LAW OF LARGE NUMBERS1

ANATOLE BECK

Introduction. The strong law of large numbers can be shown under

Theorem. Let \( X \) be a Banach space and let \( \{X_i\} \) be a sequence of
independent random \( X \)-variables (see definition below) with \( E(X_i) = 0 \),
all \( i > 0 \). Under appropriate conditions on \( X \) and on \( \{X_i\} \), we can then
assert that \( (1/n) \sum_{i=1}^{n} X_i \) converges to 0 in the strong topology of \( X \)
almost surely.

In a recent paper [1], this author showed this theorem under the
hypotheses that \( X \) is uniformly convex and that the variances of \( X_i \)
are uniformly bounded \( \text{Var}(X_i) = E(\|X_i\|^2) \). At the same time, an
example was given of a space in which the theorem fails. It is now
possible, using the methods of [1], to show a necessary and sufficient
condition on the Banach space \( X \) to yield this particular strong law
of large numbers.

A Banach space \( X \) is said to have property (A) if, for every sequence
\( \{X_i\} \) of independent random \( X \)-variables with \( E(X_i) = 0 \), all \( i \), and
\( \text{Var}(X_i) < M \), all \( i \), we have

\[
\frac{1}{n} \sum_{i=1}^{n} X_i \to 0 \quad \text{strongly almost surely.}
\]

A Banach space \( X \) is said to have property (B) if there exists an
integer \( k > 0 \) and an \( \epsilon > 0 \) such that any choice \( a_1, a_2, \ldots, a_k \) of ele-
mements from \( X \) with \( \|a_i\| \leq 1 \) gives us

\[
\| \pm a_1 \pm a_2 \pm \cdots \pm a_k \| < k(1 - \epsilon)
\]

for some combination of the + and − signs.

We shall show that these two conditions are equivalent.

1. Definitions. Let \( X \) be a separable Banach space and let

\footnote{Presented to the Society, January 23, 1961; received by the editors February 14, 1961.}

\footnote{This research was supported by Cornell University under contract with the
Office of Naval Research and by the University of Wisconsin under contract number
AF 49(638)-868 with the Air Force Office of Scientific Research.}

\footnote{Here and hereafter in this paper unless otherwise specified, all limits are taken
as \( n \to \infty \).}
(S, Σ, m) be a measure space. Then A mapping X from S into ℳ is called strongly measurable if $X^{-1}(B) = \{ s | X(s) \in B \}$ is measurable for every Borel set $B \subseteq ℳ$. If X is strongly measurable and if $\int_B \| X(s) \| m ds < \infty$ then it can be shown that there is a $y \in ℳ$ such that $x^*(y) = \int_{X(s)} x^*(X(s)) m ds$ for every $x^* \in ℳ^*$. y is defined as the integral of X.

A probability space (customarily denoted $(\Omega, B, Pr)$) is a measure space of total measure 1. (Pr(Ω) = 1.) A strongly measurable function X from Ω into ℳ is called a random ℳ-variable, and its integral, if it has one, is called its expectation, $E(X)$. If $X_1, \ldots, X_m$ are random ℳ-variables and if for every choice $B_1, \ldots, B_m$ of Borel sets from ℳ, we have

$$\Pr\{ X_1 \in B_1, \ldots, X_m \in B_m \} = \prod_{i=1}^m \Pr(X_i \in B_i),$$

then $X_1, \ldots, X_m$ are an independent collection of random ℳ-variables. If, in an infinite collection $\{ X_\alpha, \alpha \in A \}$ of random ℳ-variables, every finite sub-collection is independent, then the infinite collection is said to be independent.

For each random ℳ-variable X, we define $\text{Var}(X) = E(\|X - E(X)\|^2) = \int_\Omega \|X(\omega) - E(X)\|^2 Pr d\omega$. For each sequence $\{X_i\}$ of random ℳ-variables, we define

$$c\{X_i\} = \text{ess sup}_\Omega \limsup_n \left\| \frac{1}{n} \sum_{i=1}^n X_i \right\|.$$ 

A random ℳ-variable X is called symmetric if there is a measure-preserving mapping $\phi$ of $\Omega$ into $\Omega$ such that $X(\phi(\omega)) = -X(\omega)$ for (almost) all $\omega \in \Omega$.

2. Condition (B) implies condition (A).

THEOREM 1. Let ℳ be a Banach space satisfying condition (B) and let $\{X_i\}$ be a sequence of independent random ℳ-variables with $E(X_i) = 0$ and $\text{Var}(X_i) < M$, $i=1, 2, \ldots$. Then

$$\frac{1}{n} \sum_{i=1}^n X_i \to 0 \quad \text{strongly in ℳ almost surely.}$$

Instead of proving Theorem 1, we prove Lemma 2, below, which is the same result under additional hypotheses. The derivation of Theorem 1 from Lemma 2, i.e., the exorcism of the extraneous hypotheses, can be taken verbatim from [1], since the convexity condition is used only in Lemma 2.

1 I.e., for almost all $\omega \in \Omega$. 
Lemma 2. If $\mathcal{X}$ satisfies condition (B) and \{ $X_i$ \} is a sequence of random $\mathcal{X}$-variables and

1. the $X_i$ are independent,
2. $E(X_i) = 0$, $i = 1, 2, \ldots$,
3. $\|X_i(\omega)\| \leq 1$, all $i = 1, 2, \ldots$; $\omega \in \Omega$,
4. $X_i$ is symmetric, $i = 1, 2, \ldots$,

then

$$\frac{1}{n} \sum_{i=1}^{n} X_i \rightarrow 0 \text{ strongly in } \mathcal{X} \text{ almost surely.}$$

Proof. We will designate the sequences satisfying these hypotheses as being of type 2.

Then we can read the lemma as saying that if $\mathcal{X}$ satisfies condition (B) and \{ $X_i$ \} is of type 2, then $c\{ X_i \} = 0$. If we set $C = C(\mathcal{X}) = \sup(c\{ X_i \} | \{ X_i \} \text{ of type 2})$, then we are to prove that $C = 0$.

We assume, contrarily, that $C \neq 0$. (Note that $c\{ X_i \}$ always exists under these hypotheses, and does not exceed 1, so that $C$ exists and is no greater than 1.) We shall derive a contradiction. Choose any $\eta > 0$, and let \{ $U_i$ \} be chosen so that \{ $U_i$ \} is of type 2, and $c\{ U_i \} > C - \eta$. Let $k$ and $\epsilon$ be the numbers given to us in the definition of condition (B), and consider the random $\mathcal{X}$-variables

$$V_i = \frac{U_{ki} + U_{ki-1} + \cdots + U_{ki-k+1}}{k}.$$ 

It is easily seen that \{ $V_i$ \} is of type 2, and $c\{ V_i \} = c\{ U_i \}$. Furthermore, we can show that $E(\| V_i \|) < 1 - \epsilon/2^k$. Let $\phi_i$ be chosen so that

$$U_i(\phi_i(\omega)) = -U_i(\omega), \quad U_j(\phi_i(\omega)) = U_j(\omega),$$

all $\omega \in \Omega$, $i = 1, 2, \ldots$, $j \neq i$. Then if we look at the $2^k$ mappings

$${\phi_{ki}, \phi_{ki-1}, \ldots, \phi_{ki-k+1}}$$

given by the possible choices of $\alpha_j = 0, 1, j = 1, \ldots, k$, we see that all these are measure-preserving transformations on $\Omega$, and that for every $\omega \in \Omega$, some one of these, $\Phi_\omega$, has the property that

$$\left\| \sum_{j=0}^{k-1} U_j(\Phi_\omega(\omega)) \right\| = \left\| \pm U_{ki-k+1}(\omega) \pm \cdots \pm U_{ki}(\omega) \right\| < k(1 - \epsilon).$$

Therefore, if we number these $2^k$ mappings as $\Phi_1, \Phi_2, \ldots, \Phi_{2^k}$, we have

$^4$ Possibly this may require a change to an equivalent sequence of random variables in an isomorphic measure space. The truth of this lemma survives such a transplanting.
\[
\sum_{r=1}^{2^k} \left\| \sum_{j=k^{-1}-k+1}^{k^i} U_j(\Phi_r(\omega)) \right\| < k(2^k - 1) + k(1 - \epsilon) = k(2^k - \epsilon)
\]
for each \( \omega \in \Omega \), and therefore
\[
2^k E \left( \left\| \sum_{j=k^{-1}-k+1}^{k^i} U_j \right\| \right) = E \left( \sum_{r=1}^{2^k} \left\| \sum_{j=k^{-1}-k+1}^{k^i} U_j(\Phi_r(\cdot)) \right\| \right) < k(2^k - \epsilon),
\]
so that
\[
E(\| V_i \|) = \frac{1}{k} E \left( \sum_{j=k^{-1}-k+1}^{k^i} U_j \right) < \frac{1}{k} \cdot \frac{k(2^k - \epsilon)}{2^k} = 1 - \frac{\epsilon}{2^k}.
\]
Remember that \( \epsilon \) and \( k \) are constants depending only on the space \( \mathcal{X} \).
Let \( t > 1/\eta^2 \), and for each \( i > 0 \), define
\[
W_i = \frac{V_{ni} + V_{n-1} + \cdots + V_{ni-n+1}}{t}.
\]
We easily see that \( \{ W_i \} \) is of type 2, and that \( c\{ W_i \} = c\{ V_i \} \). Since \( \text{Var}(\| V_i \|) < 1 \), all \( i = 1, 2, \cdots \), and the \( V_i \) are independent, we have
\[
\text{Var} \left( \sum_{j=i-t-i+1}^{i} \frac{V_j}{t} \right) < \frac{1}{t}.
\]
Thus, for each \( i > 0 \),
\[
\Pr \left\{ \| W_i \| > 1 - \frac{\epsilon}{2^k} + \eta \right\} = \Pr \left\{ \sum_{j=i-t-i+1}^{i} \frac{V_j}{t} > 1 - \frac{\epsilon}{2^k} + \eta \right\}
\leq \Pr \left\{ \sum_{j=i-t-i+1}^{i} \left\| \frac{V_j}{t} \right\| > 1 - \frac{\epsilon}{2^k} + \eta \right\}
< \frac{1}{\eta^2} < \eta,
\]
by Chebyshev's inequality. Using the independence of the \( W_i \), we define
\[
Y_i = W_i, \quad Z_i = 0 \quad \text{if} \quad \| W_i \| \leq 1 - \frac{\epsilon}{2^k} + \eta,
\]
\[
i = 1, 2, \cdots
\]
\[
Y_i = 0, \quad Z_i = W_i \quad \text{if} \quad \| W_i \| > 1 - \frac{\epsilon}{2^k} + \eta,
\]
Then \( \{ Y_i \} \) is of type 2, and in fact \( \| Y_i(\omega) \| \leq 1 - \varepsilon/2^k + \eta \), all \( \omega \in \Omega, \ i = 1, 2, \ldots \), so that \( c \{ Y_i \} \leq C(1 - \varepsilon/2^k + \eta) \). The sequence \( \{ \| Z_i \| \} \) is independent, and since \( \| Z_i(\omega) \| \leq 1 \), all \( \omega \in \Omega, \ i = 1, 2, \ldots \) and \( \Pr \{ Z_i = 0 \} > 1 - \eta \), we have \( E(\| Z_i \|) < \eta \). Thus \( c \{ Z_i \} \leq c \{ \| Z_i \| \} \leq \eta \), by the (real-valued) strong law of large numbers. It is easily seen that \( c \{ W_i \} \leq c \{ Y_i \} + c \{ Z_i \} \), and thus

\[
C - \eta < c \{ U_i \} = c \{ V_i \} = c \{ W_i \} \\
\leq c \{ Y_i \} + c \{ Z_i \} \\
\leq C \left( 1 - \frac{\varepsilon}{2^k} + \eta \right) + \eta.
\]

Since \( C \leq 1 \), we have \( C(\varepsilon/2^k) < 3\eta \) for every \( \eta > 0 \) which is impossible if \( C > 0 \). This proves the lemma.

3. Condition (A) implies condition (B).

**Theorem 3.** Let \( \mathcal{X} \) be a Banach space in which condition (B) fails. Then there is in \( \mathcal{X} \) a sequence \( \{ X_i \} \) of type 2 such that

\[
c \{ X_i \} = 1.
\]

**Proof.** If \( \mathcal{X} \) fails to meet condition (B), then for every \( k \) and \( \varepsilon \), there are \( k \) vectors \( a_1, \ldots, a_k \) in the unit ball of \( \mathcal{X} \) such that

\[
\| \pm a_1 \pm \cdots \pm a_k \| \geq k(1 - \varepsilon)
\]

for every choice of signs. We now choose any sequences \( \{ e_i \} \) and \( \{ \delta_i \} \) of positive real numbers with \( e_\infty \to 0 \) and \( \delta_\infty \to 0 \). We now proceed as follows:

We choose an integer \( k_1 \) with \( k_1 > (1 - \delta_1)/\delta_1 \) and a set of elements \( a_1^{(1)}, \ldots, a_{k_1}^{(1)} \) such that

\[
\| \pm a_1^{(1)} \pm a_2^{(1)} + \cdots + a_{k_1}^{(1)} \| \geq k_1(1 - \varepsilon_1)
\]

for all choices of signs. Then, for each \( n > 1 \), we set

\[
m_n = \sum_{i=1}^{n-1} k_i
\]

and choose

\[
k_n > \frac{1 - \delta_n}{\delta_n} m_n
\]

and \( k_n \) elements \( a_1^{(n)}, \ldots, a_{k_n}^{(n)} \) such that

\[
\| \pm a_1^{(n)} \pm a_2^{(n)} + \cdots + a_{k_n}^{(n)} \| \geq k_n(1 - \varepsilon_n)
\]

* For definition of type 2, see proof of Lemma 2.
for all choices of signs. This gives us
\[ \frac{k_n}{m_{n+1}} > 1 - \delta_n \]
and
\[ \frac{m_n}{m_{n+1}} < \delta_n. \]

For any integer \( i \), we have \( m_i < i \leq m_{i+1} \) for some value of \( j \), i.e., \( i = m_j + r \), where \( 1 \leq r \leq k_j \). Define \( b_i = a_i \). This gives us a sequence \( \{b_i\} \) of elements of \( \mathcal{X} \). We define a sequence \( \{X_i\} \) of random \( \mathcal{X} \)-variables by requiring that the \( X_i \) be independent and that \( \Pr\{X_i = b_i\} = \Pr\{X_i = -b_i\} = 1/2 \). Then \( \{X_i\} \) is a sequence of type 2, and for each \( j \) and each \( \omega \in \Omega \), we have

\[
\left\| \frac{1}{m_{j+1}} \sum_{i=1}^{m_{j+1}} X_i(\omega) \right\| \geq \left\| \frac{1}{m_{j+1}} \sum_{i=m_j+1}^{m_{j+1}} X_i(\omega) \right\| - \left\| \frac{1}{m_{j+1}} \sum_{i=1}^{m_j} X_i(\omega) \right\|
\geq \frac{1}{m_{j+1}} \left\| \pm a_1^{(j)} \pm a_2^{(j)} \pm \cdots \pm a_{k_j}^{(j)} \right\| - \frac{1}{m_{j+1}} m_j
\geq \frac{1}{m_{j+1}} k_j (1 - \epsilon_j) - \frac{m_j}{m_{j+1}}
> (1 - \delta_j)(1 - \epsilon_j) - \delta_j.
\]

Thus, for every \( \omega \in \Omega \),

\[
\limsup_n \left\| \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \right\| = 1,
\]
so that \( c \{X_i\} = 1 \), as required.

**Corollary 4.** If \( \mathcal{X} \) is a Banach space, then the constant \( C(\mathcal{X}) \), defined in the proof of Lemma 2, can take only the values 0 or 1.

**Bibliography**


**Cornell University and University of Wisconsin**