The Annals of Probability
1980, Vol. 8, No. 6, 1003–1036

ALMOST SURE INVARIANCE PRINCIPLES FOR PARTIAL SUMS OF MIXING *B*-VALUED RANDOM VARIABLES.[1]

BY J. KUELBS AND WALTER PHILIPP

University of Wisconsin and University of Illinois

The approximation of partial sums of ϕ-mixing random variables with values in a Banach space B by a B-valued Brownian motion is obtained. This result yields the compact as well as the functional law of the iterated logarithm for these sums. As an application we strengthen a uniform law of the iterated logarithm for classes of functions recently obtained by Kaufman and Philipp (1978). As byproducts we obtain necessary and sufficient conditions for an almost sure invariance principle for independent identically distributed B-valued random variables and an almost sure invariance principle for sums of d-dimensional random vectors satisfying a strong mixing condition.

1. Introduction. Many of the classical limit theorems of probability hold for sequences of real-valued random variables which are weakly dependent in one sense or other. However, for weakly dependent Banach space valued random variables very little is known. In this paper we establish a central limit theorem, several laws of the iterated logarithm, as well as an almost sure invariance principle for ϕ-mixing sequences of random variables with values in a separable Banach space.

One method of proving the law of the iterated logarithm, as well as the functional law of the iterated logarithm, for a sequence of real-valued random variables is to establish an almost sure invariance principle, i.e., an approximation of the partial sum process by a suitable Brownian motion, and then to use the behavior of the Brownian motion to obtain results for the partial sum process. This is the idea of Strassen (1964) who showed that if $\{x_j, j \geq 1\}$ is a sequence of independent identically distributed real-valued random variables centered at expectations and with variance 1, then one can redefine (if necessary) the sequence $\{x_j, j \geq 1\}$ on a new probability space, on which there exists a Brownian motion $\{X(t), t \geq 0\}$ such that with probability 1

$$\sum_{j \leq t} x_j - X(t) = o((t \log \log t)^{1/2}) \tag{1.1}$$

as $t \rightarrow \infty$.

Using Brownian motion to approximate the partial sums of weakly dependent random variables such as martingale differences, and mixing and lacunary sequences has been the theme of a considerable amount of recent research. However, the methods used in this area (martingale approximation and Skorohod embedding) have seemingly been unsuitable even for independent Banach space valued random variables and hence little progress has been made in this direction. Now using the recent approximation results of Berkes and Philipp (1979) we prove a Banach space analogue of Strassen's (1964) result while handling ϕ-mixing random variables as well.

Before stating three of our theorems we will introduce some notation. Throughout the paper B denotes a real separable Banach space with norm $\|\cdot\|$, and $a_n = (2n \log \log n)^{1/2}$ where $\log x$ stands for $\log(\max(x, e))$. If $\{x_j, j \geq 1\}$ is a sequence of B-valued random variables we define $S_n = \sum_{j \leq n} x_j$, $n \geq 1$ and let $\mathscr{M}_a^b$ be the σ-field generated by the random variables x_a,

Received November 13, 1977; revised June 15, 1979.
[1] Research supported in part by NSF grants.
AMS 1970 *subject classifications.* Primary 60F15, 60G17; secondary 60B05, 60B15, 60B10.
Key words and phrases. Brownian motion in a Banach space, mixing random variables, almost sure invariance principles, law of the iterated logarithm, approximation of partial sums of Banach space valued variables.

$x_{a+1}, \ldots, x_b$. The sequence $\{x_j, j \geq 1\}$ is ϕ-mixing if there exists a sequence of real numbers $\phi(k) \downarrow 0$ such that

$$|P(A \cap B) - P(A)P(B)| \leq \phi(k)P(A) \tag{1.2}$$

for all $A \in \mathcal{M}_1^n$, $B \in \mathcal{M}_{n+k}^\infty$ and all $k, n \geq 1$. The sequence $\{x_j, j \geq 1\}$ is weakly stationary if $E\{f(x_1)g(x_n)\} = E\{f(x_{k+1})g(x_{k+n})\}$ for all $n, k \geq 1$ and $f, g \in B^*$.

We say that the sequence $\{x_j, j \geq 1\}$ satisfies the bounded law of the iterated logarithm if

$$\limsup_{n\to\infty} \| S_n \| / a_n < \infty \qquad \text{a.s.} \tag{1.3}$$

We refer to the following refinement of (1.3) as the compact law of the iterated logarithm if there is a compact set K in B such that

$$\limsup_{n\to\infty} \| S_n / a_n - K \| = 0 \qquad \text{a.s.} \tag{1.4}$$

and

$$C(\{S_n/a_n\}) = K \qquad \text{a.s.} \tag{1.5}$$

Here $\| x - K \| = \inf_{y \in K} \| x - y \|$ and $C(\{y_n\})$ denotes the cluster set of the sequence $\{y_n\}$.

Of course, if $\{x_j, j \geq 1\}$ satisfies the compact law of the iterated logarithm, then it also satisfies the bounded law of the iterated logarithm. But the converse is not necessarily true. For an example of this see Pisier (1975).

THEOREM 1. *Let $\{x_j, j \geq 1\}$ be a weakly stationary sequence of ϕ-mixing B-valued random variables with $(2 + \delta)$th moments uniformly bounded by 1. We assume that $0 < \delta \leq 1$ and that for some $\epsilon > 0$*

$$\phi(n) \ll n^{-(2+\epsilon)(1+2/\delta)}. \tag{1.6}$$

Moreover, suppose that for every $\rho > 0$ there exists a mapping $\Lambda_\rho: B \to B$ with finite dimensional range satisfying

$$E\{\Lambda_\rho(x_j)\} = 0 \qquad j \geq 1 \tag{1.7}$$

and

$$\sup_{j\geq 1} E \| \Lambda_\rho(x_j) \|^{2+\delta} < \infty. \tag{1.8}$$

Furthermore, suppose that

$$E \| \textstyle\sum_{j=a+1}^{a+n} (x_j - \Lambda_\rho(x_j)) \|^2 \leq n\rho \tag{1.9}$$

for all $\rho > 0$, $a \geq 0$ and $n \geq 1$. Then the two series defining the covariance function T of the sequence $\{x_j, j \geq 1\}$, defined as

$$T(f, g) = E\{f(x_1)g(x_1)\} + \textstyle\sum_{j\geq 2} E\{f(x_1)g(x_j)\} + \sum_{j\geq 2} E\{f(x_j)g(x_1)\} \tag{1.10}$$

converge absolutely for all $f, g \in B^*$. Moreover, without changing its distribution we can redefine the sequence $\{x_j, j \geq 1\}$ on a new probability space on which there exists a Brownian motion $\{X(t), t \geq 0\}$ with covariance structure given by T such that with probability 1

$$\| \textstyle\sum_{j\leq t} x_j - X(t) \| = o((t \log\log t)^{1/2}) \qquad t \to \infty. \tag{1.11}$$

As usual, we use the symbol $\ll$ to denote that the left-hand side is bounded by an unspecified constant times the right-hand side; in other words, we use the $\ll$ symbol instead of the O notation.

The following result is a corollary to the proof of Theorem 1.

COROLLARY 1. *Under the hypotheses of Theorem 1 the sequence $\{n^{-1/2}S_n, n \geq 1\}$ converges weakly to a mean zero Gaussian measure μ with covariance structure given by T; thus $\mu = \mathcal{L}(X(1))$.*

As a matter of fact Corollary 1 is even true for random variables satisfying a strong mixing condition with a slower rate of decay (see Proposition 4.2).

COROLLARY 2. *Let* $\{x_j, j \geq 1\}$ *be a strict sense stationary sequence of* ϕ*-mixing random variables with values in a separable Hilbert space* H. *Suppose that* x_1 *has mean zero and finite* $(2 + \delta)$th *moment. Moreover, assume that* (1.6) *is satisfied. Then the conclusions of Theorem* 1 *and Corollary* 1 *remain valid,* i.e., *an almost sure invariance principle and a central limit theorem hold for the partial sums of the sequence.*

If the maps Λ_ρ of Theorem 1 are replaced by the linear maps Π_N described in Lemma 2.1 of Kuelbs (1976a), and (1.9) is replaced by an analogous condition which relates the dimension of the range of Π_N to the rate at which the quantity corresponding to ρ can be made small, then we can strengthen (1.11). Recall that the range of the map Π_N is a Hilbert space H_μ which is a subset of B determined by the covariance function T given in (1.10) and $\mu = \mathscr{L}(Z)$ where Z is a mean zero Gaussian random variable with covariance T. Our results in this direction are not necessarily best possible, but we include them as we are certain that the methods we use can be applied to various situations even though our theorem may not apply directly. In fact since the writing of this paper Herold Dehling (1980) has substantially improved Corollary 3 below.

THEOREM 2. *Let* $\{x_j, j \geq 1\}$ *be a weakly stationary sequence of* ϕ*-mixing* B*-valued random variables, centered at expectations and with* $(2 + \delta)$th *moments uniformly bounded by* 1. *Suppose that* $0 < \delta \leq 1$ *and that* (1.6) *holds. Let* Π_N *be the linear maps determined from the covariance function* T *as given in* (1.10) *by the method described in Lemma* 2.1 *of Kuelbs* (1976a) *and suppose that*

$$\|\Pi_N\|_1 \ll e^N, \tag{1.12}$$

where $\|\Pi_N\|_1$ *denotes the operator norm when* Π_N *is considered as an operator from* B *to* H_μ. *Furthermore, assume that there is a constant* C *such that*

$$E\|\textstyle\sum_{j=a+1}^{a+n} (x_j - \Pi_N x_j)\|^2 \leq CnN^{-12/\delta} \tag{1.13}$$

for all $a \geq 0$, n, $N \geq 1$. *Then the conclusion of Theorem* 1 *holds with* (1.11) *replaced by*

$$\|\textstyle\sum_{j\leq t} x_j - X(t)\| \ll t^{1/2}(\log t)^{-1/2} \quad \text{a.s.} \tag{1.14}$$

In case $\{x_j, j \geq 1\}$ is a strictly stationary sequence with values in a separable Hilbert space the central limit assertion of Corollary 2 aides us in formulating Theorem 2 in a more direct form. The reason for this is, essentially, that the maps Π_N, determined as in Theorem 2 from the covariance function T in (1.10) or equivalently from the limiting Gaussian measure $\mu = \mathscr{L}(X(1))$ of Corollary 2, are such that we can obtain an upper bound for $\|\Pi_N\|_1$ directly and such that they converge pointwise to the identity map on the support of μ. Of course, by the support of any probability measure ν we mean the set

$$F = \{x \in H: \nu(U) > 0 \quad \text{for all open sets } U \text{ containing } x\}.$$

Further, it is known that the support of any mean zero Gaussian measure is a closed subspace. In fact, it is well-known that if Z is a H-valued random variable with distribution μ, then there is an orthonormal set $\{e_i, i \geq 1\}$ (not necessarily complete) in H such that with probability 1

$$Z = \textstyle\sum_{i\geq 1} (Z, e_i)e_i \tag{1.15}$$

where $\{Z, e_i), i \geq 1\}$ are independent mean zero Gaussian random variables with positive variances. Moreover, the support of μ is the closed subspace M generated by $\{e_i, i \geq 1\}$. These results are contained in a paper by Jain and Kallianpur (1970). In view of (1.15) and Lemma 2.1 of Kuelbs (1976a) one now can easily check that if $\lambda_i = E(Z, e_i)^2$, $i \geq 1$, then

$$\Pi_N(x) = \textstyle\sum_{i\leq N} (x, e_i)e_i = \sum_{i\leq N} (x, \alpha_i)S\alpha_i \qquad x \in H \tag{1.16}$$

where $\alpha_i = \lambda_i^{-1/2} e_i$ and $S\alpha_i = \lambda_i^{1/2} e_i$ where S is the operator of Lemma 2.1 of Kuelbs. The point in writing both expressions for Π_N in (1.16) is that both are useful in the proof of Corollary 3. We also recall for the reader that $\{S\alpha_i : i \geq 1\}$ is orthonormal in H_μ since $E(Z, \alpha_i)^2 = 1$ and $\{(Z, \alpha_i) : i \geq 1\}$ are independent. Moreover, if L denotes the projection of H onto the orthogonal complement of M we have

$$x_1 = \textstyle\sum_{i \geq 1} (x_1, e_i) e_i + L(x_1). \tag{1.17}$$

We now can formulate Corollary 3.

COROLLARY 3. *Let* $\{x_j, j \geq 1\}$ *be a strict sense stationary sequence of* ϕ*-mixing random variables with values in a separable Hilbert space* H. *Assume that* x_1 *has mean zero and finite* $(2 + \delta)$th *moment with* $0 < \delta \leq 1$. *Suppose that* (1.6) *holds and that if* x_1 *is written in the form* (1.17), *then*

$$\textstyle\sum_{i \geq N} E(x_1, e_i)^2 \ll N^{-12/\delta} \qquad \textit{and} \quad \inf_{1 \leq i \leq N} E(Z, e_i)^2 \gg e^{-N} \tag{1.18}$$

where Z *is the mean zero Gaussian random variable determined by* T *as given in* (1.10). *Then the conclusion of Theorem* 1 *holds with* (1.11) *replaced by* (1.14).

For independent random variables we also have the following theorem.

THEOREM 3. *Let* $\{x_j, j \geq 1\}$ *be independent identically distributed* B*-valued random variables centered at expectations and with finite* $(2 + \delta)$th *moments with* $0 < \delta \leq 1$. *Then the following two statements are equivalent.*

(a) *Without changing its distribution we can redefine the sequence* $\{x_j, j \geq 1\}$ *on a new probability space on which there exists a Brownian motion* $\{X(t), t \geq 0\}$ *with covariance structure given by*

$$T(f, g) = E\{f(x_1)g(x_1)\} \qquad f, g \in B^*,$$

such that with probability 1

$$\|\textstyle\sum_{j \leq t} x_j - X(t)\| = o((t \log\log t)^{1/2}) \qquad t \to \infty.$$

(b) x_1 *is pre-Gaussian and any of the following three conditions holds.*

(i) $S_n/a_n \to 0$ *in probability*

(ii) $E\|S_n\|/a_n \to 0$

(iii) $\{x_j, j \geq 1\}$ *satisfies the compact law of the iterated logarithm with limit set* K, *the unit ball of the Hilbert space* $H_{\mathscr{L}(x_1)}$.

The Hilbert space $H_{\mathscr{L}(x_1)}$ and the limit set K in Theorem 3 are defined in Lemma 2.1 of Kuelbs (1976a).

Recall that a random variable x is called pre-Gaussian if its covariance function $T(f, g) = E\{f(x)g(x)\}$ is the covariance function of a countably additive Gaussian measure on the Banach space B. For the sake of historical accuracy, however, we would like to observe that pre-Gaussian random variables are in fact post-Gaussian.

In the meantime Philipp (1979) has proved Theorem 3 assuming only finite second moments thereby generalizing Strassen's (1964) Theorem 2 to B-valued random variables.

Finally we point out the rather surprising fact that for certain independent identically distributed sequences of Banach space valued random variables, the compact law of the iterated logarithm holds, yet it is impossible to approximate the partial sum process by a Brownian motion with an error term which is $o((t \log\log t)^{1/2})$ as $t \to \infty$. This contrasts sharply with the finite dimensional situation. We will elaborate on this point in Section 5.

As we mentioned earlier, the asymptotic fluctuation behavior of Brownian motion translates immediately to any sequence of random variables for which an almost sure invariance principle holds with a sufficiently small error term. Indeed, let $\{x_j, j \geq 1\}$ be any sequence of random variables satisfying (1.11). Then $\{x_j, j \geq 1\}$ satisfies the compact law of the iterated logarithm since by Theorem 4.1 of Kuelbs (1977) the sequence $\{X(j) - X(j-1), j \geq 1\}$ does. Moreover, since in Theorem 1 $\mu = \mathscr{L}(X(1))$ the limit set K is the unit ball of the Hilbert space H_μ which generates μ.

But we also have a functional law of the iterated logarithm analogous to the real-valued case studied by Strassen (1964). Let $C_B[0, 1]$ denote the Banach space of B-valued continuous functions on $[0, 1]$ which vanish at zero, with the norm

$$\|f\|_{B,\infty} = \sup_{0 \leq t \leq 1} \|f(t)\|, \qquad f \in C_B[0, 1].$$

For $n \geq 1$ and $0 \leq t \leq 1$ define $Y_n(t) = X(nt)$ where $\{X(t), t \geq 0\}$ is Brownian motion. Then $Y_n \in C_B[0, 1]$ for $n \geq 1$. Moreover, by Theorem 1 of Kuelbs and LePage (1973) there is a compact set $K \subseteq C_B([0, 1]$ uniquely determined by the covariance of $\mathscr{L}(X(1))$ such that with probability 1

$$\lim_{n\to\infty} \| Y_n/a_n - K\|_{B,\infty} = 0 \tag{1.19}$$

and

$$C(\{Y_n/a_n, n \geq 1\}) = K. \tag{1.20}$$

The cluster set K is described in Kuelbs and LePage (1973), but also can be computed from the covariance function of the Gaussian measure P induced on $C_B[0, 1]$ by the process $\{X(t), 0 \leq t \leq 1\}$ by using Lemma 2.1 of Kuelbs (1976a).

Now let $\{x_j, j \geq 1\}$ be any sequence of random variables satisfying (1.11). For $n \geq 1$ and $0 \leq t \leq 1$ define

$$T_n(t) = S_{nt} \quad \text{if} \quad t = k/n, \quad k = 0, 1, \cdots, n$$

and linear in between these points. Then by (1.11)

$$\| Y_n - T_n \|_{B,\infty}/a_n = o(1)$$

and hence by (1.19) and (1.20) the partial sum process also satisfies the functional law of the iterated logarithm, i.e., (1.19) and (1.20) continue to hold with probability 1 if we replace Y_n by T_n.

As a final application of translating asymptotic fluctuation behavior for Brownian motion to the related partial sum process we present a corollary which combines Corollary 3 and Theorem 2.4 of Kuelbs (1975b).

COROLLARY 4. *Let $\{x_j, j \geq 1\}$ be a strict sense stationary sequence of ϕ-mixing random variables with values in a separable Hilbert space H. Assume x_1 has mean zero and finite* $(2 + \delta)$th *moment with* $0 < \delta \leq 1$. *Suppose that* (1.6) *and* (1.18) *hold. Let Z be as in* (1.15), *set*

$$\Gamma = \sup_{i \geq 1} \{E(Z, e_i)^2\}^{1/2}, \tag{1.21}$$

and let n_1 denote the cardinality of the set of those integers i where the supremum is achieved in (1.21). *Finally, let $\psi(t)$ be a positive nondecreasing, continuous function defined on* $[1, \infty)$. *Then*

$$P\{\| S_n\| > n^{1/2}\psi(n)\Gamma \text{ i.o.}\} = 0 \quad \textit{or} \quad 1 \tag{1.22}$$

according as

$$\int_1^\infty \frac{(\psi(t))^{n_1}}{t} \exp(-\psi^2(t)/2)\, dt < \infty \quad \textit{or} \quad = \infty. \tag{1.23}$$

Comparing the hypotheses of Theorems 1 and 3 we notice that in Theorem 3 no assumptions are made to the effect that the random variables are approximable by finite dimensional ones. This is in direct contrast to Theorem 1 and thus (1.9) might appear overly restrictive. Indeed, (1.9) is undoubtedly not a necessary condition for (1.11). But there are many situations when the approximation in (1.9) is automatic. One example is furnished by Corollary 2. Another example occurs when $\{x_j, j \geq 1\}$ is a sequence of independent identically distributed random variables centered at expectations, with finite second moments and satisfying the central limit theorem in B. This can be seen by using the maps Π_N defined in Lemma 2.1 of Kuelbs (1976a) along with Theorem 5.1 of de Acosta and Giné (1979). A third example is provided if $\{x_j, j \geq 1\}$ is a sequence of independent identically distributed random variables with mean zero, finite second moments and assuming values in a type 2 Banach space. Recall that a Banach space is of type 2 if there is a constant A such that

$$E\|z_1 + \cdots + z_n\|^2 \leq A \sum_{j \leq n} E\|z_j\|^2$$

for all sequences $\{z_j, j \geq 1\}$ of independent mean zero random variables and all $n \geq 1$. Suppose now that there is a collection of maps $\{\Lambda_\rho, \rho > 0\}$ taking B into B with finite dimensional range and such that

$$E\{\Lambda_\rho x_1\} = 0, \qquad E\|x_1 - \Lambda_\rho x_1\|^2 \leq \rho/A.$$

Such a family of maps Λ_ρ exists by Kuelbs (1977), page 790. Then obviously (1.9) follows. Hence in this setting the approximation of a single random variable by a finite dimensional one translates immediately into the approximation of the partial sums.

We also would like to observe that if B is a Hilbert space or even only a type 2 Banach space the rate of decay for ϕ demanded in (1.6) can be considerably relaxed.

The reader may wonder why we have Theorem 1 only for ϕ-mixing random variables and not for random variables satisfying a strong mixing condition. The main reason is the lack of a proper generalization of Theorem 2 of Berkes and Philipp (1979) to strong mixing random variables. But since we and Berkes are confident that such a generalization will be forthcoming in the not too distant future, we have arranged the proof in such a way that the full strength of the ϕ-mixing condition is employed only in Section 3. In Sections 2 and 4, which contain the remaining material on which the proof of Theorem 1 rests, we only assume that the random variables satisfy a strong mixing condition with an even slower rate of decay for the mixing coefficient.

As a by-product we obtain in Section 2 an almost sure invariance principle for sums of d-dimensional random vectors satisfying a strong mixing condition. In Section 3 we prove a bounded law of the iterated logarithm for ϕ-mixing B-valued random variables assuming only a rather mild growth condition on the variance of the partial sums. In this section we also have estimates on the $(2 + \delta)$th moments of and exponential bounds for the partial sums of such random variables which may be useful in other applications. In Section 4 we complete the proof of Theorems 1 and 2. We also use some of the material of Section 4 for the proof of Theorem 3, carried out in Section 5. In Section 5 we also give an example which is instructive in connection with the proof of Theorem 1. As an application of Theorem 1 we prove in Section 6 a refinement of a uniform law of the iterated logarithm recently obtained by Kaufman and Philipp (1978).

The proof of Theorem 1 is probably easiest understood by starting to read Section 4.4 and working one's way backwards.

2. An almost sure invariance principle for sums of mixing random vectors. As a by-product of the proof of the almost sure invariance principle for B-space valued random variables we obtain an almost sure invariance principle with a sharper error term for sums of random vectors satisfying a strong mixing condition. Let $\{\xi_n, n \geq 1\}$ be a sequence of random vectors $\in \mathbb{R}^d$ and let $\mathcal{M}_a^b$ be the σ-field generated by the random vectors $\xi_a, \xi_{a+1}, \ldots, \xi_b$. Then $\{\xi_n, n \geq 1\}$ is said to satisfy a strong mixing condition if there exists a nonincreasing sequence $\{\rho(n), n \geq 1\}$ such that

$$|P(AB) - P(A)P(B)| \leq \rho(n) \downarrow 0 \tag{2.1}$$

for all $n, k \geq 1$, all $A \in \mathcal{M}_1^k$ and $B \in \mathcal{M}_{k+n}^\infty$.

THEOREM 4. *Let $\{\xi_n, n \geq 1\}$ be a weak sense stationary sequence of $\mathbb{R}^d$-valued random vectors, centered at expectations and having $(2 + \delta)$th moments with $0 < \delta \leq 1$, uniformly bounded by 1. Suppose that $\{\xi_n, n \geq 1\}$ satisfies a strong mixing condition (2.1) with*

(2.2) $$\rho(n) \ll n^{-(1+\epsilon)(1+2/\delta)} \qquad \epsilon > 0.$$

Write

$$\xi_n = (\xi_{n1}, \ldots, \xi_{nd}).$$

Then the two series in

(2.3) $$\gamma_{ij} = E\xi_{1i}\xi_{1j} + \sum_{k\ge 2} E\xi_{1i}\xi_{kj} + \sum_{k\ge 2} E\xi_{ki}\xi_{1j}$$

converge absolutely. Denote the matrix $((\gamma_{ij}))(1 \le i, j \le d)$ *by* Γ. *Then we can redefine the sequence* $\{\xi_n, n \ge 1\}$ *on a new probability space together with Brownian motion* $X(t)$ *with covariance matrix* Γ *such that*

$$\sum_{n\le t} \xi_n - X(t) \ll t^{1/2-\lambda} \qquad \text{a.s.}$$

for some $\lambda > 0$ *depending on* ϵ, δ *and* d *only.*

For independent random vectors $\in \mathbb{R}^d$ this result was recently proved by Berkes and Philipp (1979). Moreover, in that paper a sketch of the proof of Theorem 4 was given for the case $d = 1$ using the present method. (See Remark 4.4.4 of Berkes and Philipp (1979).) On the other hand for the case $d = 1$ a much more general almost sure invariance principle was proved by Philipp and Stout (1975) using martingale approximation and the Skorohod embedding theorem (see Theorems 7.1, 8.1 and 8.2 and 11.1 of Philipp and Stout (1975)). It is possible to extend Theorems 7.1, 8.1 and 8.2 of Philipp and Stout (1975) to random vectors using the present method.

Of course, by a weak sense stationary sequence of random vectors $\{\xi_n, n \ge 1\}$ we mean a sequence whose components ξ_{nj} satisfy

$$E\{\xi_{a+n,j}\xi_{a+1,i}\} = E\{\xi_{nj}\xi_{1i}\}$$

for all $a \ge 0$, $n \ge 1$ and $1 \le i, j \le d$. It is easy to see that this is in agreement with the definition for B-space valued random variables as given in Section 1.

2.1. *Preliminaries.*

LEMMA 2.1. *Let* ξ *and* η *be two random variables measurable* $\mathscr{F}$ *and* $\mathscr{G}$ *respectively. Let* $r, s, t \ge 1$ *with* $r^{-1} + s^{-1} + t^{-1} = 1$. *If* $\|\xi\|_s < \infty$ *and* $\|\eta\|_t < \infty$ *then*

(2.4) $$|E\xi\eta - E\xi E\eta| \le 10(\rho(\mathscr{F}, \mathscr{G}))^{1/r}\|\xi\|_s\|\eta\|_t.$$

Moreover, if $\|\xi\|_\infty < \infty$ *and* $\|\eta\|_\infty < \infty$ *then*

(2.5) $$|E\xi\eta - E\xi E\eta| \le 4\rho(\mathscr{F}, \mathscr{G})\,\|\xi\|_\infty\,\|\eta\|_\infty.$$

Here

$$\rho(\mathscr{F}, \mathscr{G}) = \sup|P(AB) - P(A)P(B)|$$

the supremum being extended over all $A \in \mathscr{F}$ *and* $B \in \mathscr{G}$.

Relation (2.5) is due to Volkonskii and Rozanov (1959), (2.4) is due to Davydov (1970). For a proof see Deo (1973). The next lemma is due to Dvoretzky (1970).

LEMMA 2.2. *Let* ξ *be a (possibly complex-valued) random variable with* $|\xi| \le 1$ *and let* $\mathscr{F}$ *be the* σ*-field generated by* ξ. *Then for any* σ*-field* $\mathscr{G}$

$$E|E(\xi|\mathscr{G}) - E\xi| \le 2\pi\rho(\mathscr{F}, \mathscr{G}).$$

If ξ *is real-valued then the constant* 2π *can be replaced by* 4.

This lemma is implicitly contained in the proof of (2.5) as given in Ibragimov and Linnik (1971), page 306. On the other hand if we take the well-known relation (2.5) for granted we can give a very simple proof in the real case. Assume without loss of generality $E\xi = 0$. Then by (2.5)

$$E|E(\xi|\mathscr{G})| = E\{E(\xi|\mathscr{G})\cdot\text{sign } E(\xi|\mathscr{G})\}$$
$$= E\{\xi\cdot\text{sign } E(\xi|\mathscr{G})\} \le 4\rho(\mathscr{F}, \mathscr{G}).$$

We observe that similar results, such as Theorem 2.2 of Serfling (1968) can be proved in the same way by combining this trick with one used in Philipp and Stout (1975), page 33.

In Section 4 we shall need to allow that the dimension d tends to infinity at a certain rate. For this reason it is important to keep track of how the estimates in the following lemmas depend on d.

From now on we assume that

$$0 < \epsilon \le \tfrac{1}{4}. \tag{2.6}$$

LEMMA 2.3. *Let $\{\xi_n, n \ge 1\}$ be as in Theorem 4. Then for all $1 \le i, j \le d$*

$$E\{\textstyle\sum_{k,m\le n} \xi_{ki}\xi_{mj}\} = \gamma_{ij}n + O(n^{1-\epsilon})$$

and

$$\gamma_{ij} = O(1) \tag{2.7}$$

where the constants implied by O only depend on ϵ and the constant implied by $\ll$ in (2.2).

The proof follows along well-known lines. We have by (2.4) and (2.2)

$$|E\xi_{1i}\xi_{kj}| \le 10\|\xi_{1i}\|_{2+\delta}\|\xi_{kj}\|_{2+\delta}(\rho(k-1))^{\delta/(2+\delta)} \ll k^{-1-\epsilon}. \tag{2.8}$$

Thus

$$\begin{aligned} E\{\textstyle\sum_{k,m\le n} \xi_{ki}\xi_{mj}\} &= nE\xi_{1i}\xi_{1j} + \textstyle\sum_{k=2}^{n}(n-k+1)E\xi_{1i}\xi_{kj} \\ &\quad + \textstyle\sum_{k=2}^{n}(n-k+1)E\xi_{ki}\xi_{1j} \\ &= n(E\xi_{1i}\xi_{1j} + \textstyle\sum_{k=2}^{n}E\xi_{1i}\xi_{kj} \\ &\quad + \textstyle\sum_{k=2}^{n}E\xi_{ki}\xi_{1j}) + O(\textstyle\sum_{k=2}^{n}k^{-\epsilon}) \\ &= n\gamma_{ij} + O(n\textstyle\sum_{k>n}k^{-1-\epsilon}) + O(n^{1-\epsilon}) \\ &= n\gamma_{ij} + O(n^{1-\epsilon}). \end{aligned}$$

Relation (2.7) now follows from (2.3) and (2.8).

LEMMA 2.4. *Let $\{\eta_n, n \ge 1\}$ be a weak sense stationary sequence of random variables, centered at expectations and with $(2+\delta)$th moments uniformly bounded by 1. Suppose that $\{\eta_n, n \ge 1\}$ satisfies a strong mixing condition* (2.1) *with rate of decay given by* (2.2). *Then*

$$E(\textstyle\sum_{k\le n}\eta_k)^2 = \sigma^2 n + O(n^{1-\epsilon})$$

where

$$\sigma^2 = E\eta_1^2 + 2\textstyle\sum_{k\ge 2}E\eta_1\eta_k = O(1) \tag{2.9}$$

and where the constants implied by O only depend on ϵ and the constant implied by $\ll$ in (2.2). *We call σ^2 the variance of the sequence $\{\eta_n, n \ge 1\}$.*

This follows from Lemma 2.3 by setting $d = 1$ and $\eta_k = \xi_k$.

The next lemma is due to Sotres and Malay Ghosh (1977). The value of α can be obtained by a careful analysis of their proof and of Serfling's (1968) paper on which their proof rests.

LEMMA 2.5. *Let $\{\eta_n, n \geq 1\}$ be a sequence of random variables, centered at expectations with* $(2 + \delta)$th *moments with* $0 < \delta \leq 1$, *uniformly bounded by* 1. *Suppose that $\eta_n, n \geq 1\}$ satisfies a strong mixing condition with rate of decay $\rho(n)$ given by* (2.2). *Put $\alpha = \epsilon\delta/8$. Then for all $a \geq 0$*

$$E \mid \sum_{\nu=a+1}^{a+n} \eta_\nu \mid^{2+\alpha} \ll n^{1+\alpha/2}.$$

Here the constant implied by $\ll$ only depends on ϵ, δ and the constant implied by $\ll$ in (2.2).

2.2. *The central limit theorem.* Let $u = (u_1, \cdots, u_d) \in \mathbb{R}^d$ and let $\{\xi_n, n \geq 1\}$ be the sequence of Theorem 4. We write for $n \geq 1$

(2.10) $$\zeta_n = \langle u, \xi_n \rangle = \sum_{j \leq d} u_j \xi_{nj}.$$

Then $\{\zeta_n, n \geq 1\}$ satisfies the same mixing condition as $\{\xi_n, n \geq 1\}$.

LEMMA 2.6. *We have for the variance σ^2 of the sequence $\{\zeta_n, n \geq 1\}$*

$$\sigma^2 = \langle u, \Gamma u \rangle \ll d \mid u \mid^2$$

where the constant implied by $\ll$ only depends on ϵ and the constant implied by $\ll$ in (2.2).

This follows at once from (2.9), (2.3) and (2.7).

For the proof of Theorem 1 we need to assume that the dimension d and the bound on the $(2 + \delta)$th moments of the random vectors $\xi_n (1 \leq n \leq N)$ may increase with N. For that purpose we fix $N \geq 1$ and assume that $\{\xi_n, 1 \leq n \leq N\}$ is a sequence of random vectors of dimension $d(N)$. Let $f_N(u)$ be the characteristic function of $N^{-1/2} \sum_{n \leq N} \xi_n$.

PROPOSITION 2.1. *Let $\{\xi_n, 1 \leq n \leq N\}$ be a weak sense stationary sequence of random vectors with values in $\mathbb{R}^d$, $d = d(N)$, centered at expectations and with* $(2 + \delta)$th *moments with $0 < \delta \leq 1$, bounded by $b = b(N) \geq 1$, say. Suppose that the sequence satisfies a strong mixing condition* (2.1) *with rate of decay* (2.2). *If $d = d(N) \leq N^{\alpha/64}$ where α is given by Lemma* 2.5 *then*

$$f_N(u) - \exp(-\tfrac{1}{2} \langle u, \Gamma u \rangle) \ll bN^{-\alpha/32}$$

for all $u \in \mathbb{R}^d$ with $\mid u \mid \leq N^{\alpha/64}$. Here the constant implied by $\ll$ only depends on ϵ, δ, and the constant implied by $\ll$ in (2.2).

The proof follows the usual pattern of defining large block sums and small block sums of random variables and of estimating their characteristic functions. Since we expand the characteristic functions using a Taylor expansion up to the $(2 + \delta)$th moments we can assume without loss of generality that $b = 1$. Let

(2.11) $$p = [N^{3/4}], \; q = [N^{(1+\alpha)/4}], \qquad l = [N/(p + q)].$$

Then

(2.12) $$l = N^{1/4}(1 + O(N^{-1/4})).$$

We introduce the intervals

$$H_j = ((j - 1)(p + q), jp + (j - 1)q] \qquad 1 \leq j \leq l$$

$$I_j = (jp + (j - 1)q, j(p + q)]. \qquad 1 \leq j \leq l$$

$$I_{l+1} = (l(p + q), N]$$

and define the random variables $y_j (1 \leq j \leq l)$ and $z_j (1 \leq j \leq l + 1)$ by

$$y_j = \sum_{\nu \in H_j} \zeta_\nu, \qquad z_j = \sum_{\nu \in I_j} \zeta_\nu.$$

LEMMA 2.7. *Under the hypotheses of Proposition* 2.1

$$E \mid \sum_{j\le l+1} z_j \mid^2 \ll N^{7/8} b.$$

PROOF. Replacing ξ_n by $\xi_n b^{-1/(2+\delta)}$ we can assume $b = 1$ without loss of generality. By (2.10) and Cauchy's inequality

$$\mid \sum_{\nu=a+1}^{a+n} \zeta_\nu \mid \le \mid u \mid \mid \sum_{\nu=a+1}^{a+n} \xi_\nu \mid \le N^{\alpha/64} \sum_{j\le d} \mid \sum_{\nu=a+1}^{a+n} \xi_{\nu j} \mid$$

for all $a \ge 0$, $n \ge 1$ and $\mid u \mid \le N^{\alpha/64}$. We apply Lemma 2.5 to each sequence $\{\xi_{\nu j}, \nu \ge 1\}(1 \le j \le d)$ and obtain

$$E \mid \sum_{\nu=a+1}^{a+n} \xi_{\nu j} \mid^{2+\gamma} \ll n^{1+\gamma/2}$$

uniformly for $a \ge 0$, $1 \le j \le d$ and $0 \le \gamma \le \alpha$. More precisely the constant implied by $\ll$ only depends on ϵ, δ and the constant implied by $\ll$ in (2.2). Consequently

$$E \mid \sum_{\nu=a+1}^{a+n} \zeta_\nu \mid^{2+\gamma} \ll N^{(2+\gamma)\alpha/32} n^{1+\gamma/2} \tag{2.13}$$

with the same provision for the constant implied by $\ll$. Thus

$$E z_j^2 \ll N^{1/4+3\alpha/8} \tag{2.14}$$

and

$$E z_{l+1}^2 \ll N^{3/4+\alpha/8}. \tag{2.15}$$

Hence by Minkowski's inequality, (2.12), (2.14), (2.15) and since $\alpha \le 1/4$ by (2.6) and Lemma 2.5

$$\| \sum_{j\le l+1} z_j \|_2 \ll l \cdot N^{1/8+3\alpha/16} + N^{3/8+\alpha/16} \ll N^{3/8+3\alpha/16} \ll N^{7/16}.$$

LEMMA 2.8. *Under the hypotheses of Proposition* 2.1

$$\prod_{j\le l} E\{\exp(i y_j N^{-1/2})\} - \exp(-\tfrac{1}{2} \langle u, \Gamma u\rangle) \ll N^{-\alpha/32} b.$$

The proof is a routine modification of the standard proof of the central limit theorem for independent random variables. Again we assume $b = 1$. We first observe that by Lemma 2.6

$$\sigma^2 \ll N^{\alpha/16}. \tag{2.16}$$

Thus by Lemma 2.4

$$\begin{aligned} E y_j^2 &= \sigma^2 N^{3/4} + O(N^{3(1-\alpha)/4} N^{\alpha/16}) \\ &= \sigma^2 N^{3/4} + O(N^{3/4-\alpha/2}) \end{aligned} \tag{2.17}$$

since $\alpha \le \epsilon$ and since $\zeta_\nu N^{-\alpha/32}$ has $(2 + \delta)$th moment bounded by 1. Moreover, by (2.13)

$$E \mid y_j \mid^{2+\alpha} \ll N^{(2+\alpha)\alpha/32} N^{3(2+\alpha)/8} \ll N^{(1+\alpha/2)(3/4+\alpha/16)}. \tag{2.18}$$

Thus for some θ with $\mid \theta \mid \le 1$

$$E\{\exp(i y_j N^{-1/2})\} = 1 - \tfrac{1}{2} N^{-1} E y_j^2 + \theta N^{-1-\alpha/2} E \mid y_j \mid^{2+\alpha} = 1 - r_j, \quad \text{say}. \tag{2.19}$$

Now by (2.17) and (2.18)

$$\begin{aligned} r_j &= \tfrac{1}{2} N^{-1/4}(\sigma^2 + O(N^{-\alpha/2})) + O(N^{-(1+\alpha/2)(1/4-\alpha/16)}) \\ &= \tfrac{1}{2} N^{-1/4} \sigma^2 + O(N^{-1/4-\alpha/32}). \end{aligned}$$

Thus by (2.16)

$$r_j^2 \ll N^{-1/2+\alpha/4} \ll N^{-3/8}.$$

Consequently,

$$\log E\{\exp(iy_j N^{-1/2}) = -r_j + \theta r_j^2 = -\tfrac{1}{2} N^{-1/4}\sigma^2 + O(N^{-1/4-\alpha/32})$$

and

$$\log \prod_{j\le l} E\{\exp(iy_j N^{-1/2}) = -\tfrac{1}{2}\sigma^2(1 + O(N^{-1/8})) + O(N^{-\alpha/32}) = -\tfrac{1}{2}\sigma^2 + O(N^{-\alpha/32})$$

by (2.12) and (2.16). Hence

$$\prod_{j\le l} E\{\exp(iy_j N^{-1/2})\} = e^{-\sigma^2/2} e^{O(N^{-\alpha/32})}) = e^{-\sigma^2/2}(1 + O(N^{-\alpha/32})).$$

The lemma follows now from Lemma 2.6. □

We can now finish the proof of Proposition 2.1 in the usual fashion. We have

$$f_N(u) - \exp(-\tfrac{1}{2}\langle u, \Gamma u\rangle) = I + II + III, \quad \text{say,} \tag{2.20}$$

where

$$I = E\{\exp(iN^{-1/2}(\textstyle\sum_{j\le l} y_j + \sum_{j\le l+1} z_j))\} - E\{\exp(iN^{-1/2}\sum_{j\le l} y_j)\}$$

$$II = E\{\exp(iN^{-1/2}\textstyle\sum_{j\le l} y_j)\} - \prod_{j\le l} E\{\exp(iN^{-1/2} y_j)\}$$

$$III = \textstyle\prod_{j\le l} E\{\exp(iN^{-1/2} y_j)\} - \exp(-\tfrac{1}{2}\langle u, \Gamma u\rangle).$$

Since for real x and y

$$|e^{i(x+y)} - e^{iy}| = |e^{ix} - 1| \le |x|$$

we have by Lemma 2.7 and Cauchy's inequality

$$I \ll N^{-1/2} E|\textstyle\sum_{j\le l+1} z_j| \ll N^{-1/2} N^{7/16} b \ll N^{-1/16} b.$$

We apply Lemma 2.1 l times to the real and imaginary parts and obtain by (2.12), (2.11) and (2.2)

$$II \ll l\rho(q) \ll N^{1/4} N^{-(1+\alpha)/4} \ll N^{-\alpha/4}.$$

Since III is estimated in Lemma 2.8 we obtain Proposition 2.1 from (2.20) and these estimates.

2.3. *Introduction of the blocks.* In Sections 4 and 5 we need to assume that the dimension d of the random variables ξ_n increases with n. However, treating the cases $d(N) = \text{const}$ and $d(N) \uparrow \infty$ simultaneously, as was done in the proof of the central limit theorem, would yield a worse error term in Theorem 4. Hence we shall assume throughout the remainder of this section that the dimension d is fixed.

We define blocks H_k and I_k of consecutive positive intergers leaving no gaps between blocks. The order is $H_1, I_1, H_2, I_2, \cdots$ (These blocks have nothing in common with the blocks introduced in Section 2.2.) To define the length of the blocks we write

$$\beta = 288\, d\alpha^{-1} \tag{2.21}$$

and set

$$\text{card } H_k = [k^\beta], \text{ card } I_k = [k^{\beta/4}] \qquad k \ge 1. \tag{2.22}$$

Put

$$t_k = \textstyle\sum_{j\le k} \text{card}(H_j \cup I_j). \tag{2.23}$$

Then

$$k^{\beta+1} \ll t_k \ll k^{\beta+1}. \tag{2.24}$$

PROPOSITION 2.2. *There is a constant* $\lambda > 0$ *such that as* $k \to \infty$

$$\max_{t_k < N \le t_{k+1}} \left| \sum_{\nu=t_k+1}^{N} \xi_\nu \right| \ll t_k^{1/2-\lambda} \qquad \text{a.s.}$$

It is enough to prove the inequality for each component $\xi_{\nu j}(1 \le j \le d)$. Denote them generically by η_ν. We put

$$F(r, s) = \left| \sum_{\nu=t_k+1+r}^{t_k+r+s} \eta_\nu \right|. \tag{2.25}$$

For given N with $t_k < N \le t_{k+1}$ we let $n = n(N)$ be the largest integer such that

$$2^n \le N - t_k. \tag{2.26}$$

Writing $N - t_k$ in dyadic expansion we have

$$N - t_k = \sum_{0 \le l \le n} \epsilon_l 2^l$$

where $\epsilon_l = 0, 1$. Hence (for the details see, e.g., S. and L. Gaal (1964), page 139)

$$F(0, N - t_k) \le \sum_{0 \le l \le n} F(m_l 2^{l+1}, 2^l) \tag{2.27}$$

where $0 \le m_l < 2^{n-l}$ $(0 \le l \le n)$ are integers. Put $\gamma = \alpha/8(\beta + 1)$ and define the events

$$G_k(m, l) = \{F(m2^{l+1}, 2^l) \ge t_k^{(1-\gamma)/2}\} \tag{2.28}$$

$$G_k = \bigcup_{l \le n_k} \bigcup_{m < 2^{n_k - l}} G_k(m, l) \tag{2.29}$$

where $n_k = n(t_{k+1})$ in the above notation.

LEMMA 2.9. *With probability* 1 *only finitely many of the events* G_k *occur.*

PROOF. We have by Lemma 2.5 and (2.24)

$$P\{G_k(m, l)\} \ll k^{-(\beta+1)(1-\gamma)(1+\alpha/2)} 2^{l(1+\alpha/2)}.$$

Thus by (2.22), (2.28) and (2.29)

$$P(G_k) \ll k^{-(\beta+1)(1-\gamma)(1+\alpha/2)} \sum_{l \le n_k} 2^{l(1+\alpha/2)} 2^{n_k - l}$$

$$\ll k^{-(\beta+1)(1-\gamma)(1+\alpha/2)} 2^{n_k(1+\alpha/2)} \ll k^{-1-\alpha/8}. \qquad \square$$

Proposition 2.2 follows now at once. Indeed, by (2.25) – (2.29) we have with probability 1

$$\max_{t_k < N \le t_{k+1}} F(0, N - t_k) \ll n_k \cdot t_k^{(1-\gamma)/2} \ll t_k^{1/2-\gamma/4}. \qquad \square$$

REMARK. A similar argument is used in Section 3.3 and in the proof of Proposition 4.1.

We now define random variables y_j and z_j by

$$y_j = \sum_{\nu \in H_j} \xi_\nu \qquad z_j = \sum_{\nu \in I_j} \xi_\nu. \tag{2.30}$$

(Again these have nothing to do with the ones introduced in the previous section).

LEMMA 2.10. *We have as* $M \to \infty$

$$\left| \sum_{j \le M} z_j \right| \ll t_M^{1/4} \qquad \text{a.s.}$$

PROOF. By Minkowski's inequality, Lemma 2.3 and (2.22)

$$\left\| \sum_{j \le M} z_j \right\|_2 \le \sum_{j \le M} \| z_j \|_2 \ll \sum_{j \le M} j^{\beta/8} \ll M^{1+\beta/8}.$$

The lemma follows now from Chebyshev's inequality and the Borel-Cantelli lemma since $\beta \ge 288/\alpha > 4 \cdot 288$.

PROOF OF THEOREM 4. We follow Berkes and Philipp (1979, pages 43, 44). Put

$$X_k = k^{-\beta/2} \cdot y_k$$

and let $\mathscr{F}_k$ be the σ-field generated by $X_1, \cdots, X_k$. Then by Lemma 2.2, (2.2) and (2.21)

$$E \mid E\{\exp(i\langle u, X_k\rangle \mid \mathscr{F}_{k-1})\} - E\{\exp(i\langle u, X_k\rangle)\}\mid \ll \rho(k^{\beta/4}) \ll k^{-\beta/4} \ll k^{-9d}.$$

From Proposition 2.1 we obtain using (2.21) once more

$$E\{\exp(i\langle u, X_k\rangle)\} - \exp(-\tfrac{1}{2}\langle u, \Gamma u\rangle) \ll (k^\beta)^{-\alpha/32} \ll k^{-9d}$$

for all u with $|u| \le k^{9/4} \le (k^\beta)^{\alpha/64}$. These two inequalities show that we can choose $\lambda_k = \text{const} \cdot k^{-9d}$ and $T_k = \text{const} \cdot k^{9/4}$. The remainder of the proof now follows Berkes and Philipp (1979, pages 43, 44) verbatim except that in (3.10) of that paper we have to include the estimate given by Lemma 2.10. We also note that Γ need not be positive definite to carry out the estimate of page 43 of Berkes and Philipp (1979).

3. The bounded law of the iterated logarithm for ϕ-mixing B-space valued random variables. In this section we prove the following theorem.

THEOREM 5. *Let $\{x_\nu, \nu \ge 1\}$ be a sequence of ϕ-mixing B-space valued random variables with* $(2 + \delta)$th *moments uniformly bounded. We assume that* $0 < \delta \le 1$ *and that*

$$\phi(n) \ll n^{-(2+\epsilon)(1+2/\delta)} \tag{3.1}$$

for some $0 < \epsilon \le \tfrac{1}{4}$. *Suppose that for some* σ

$$E \| \textstyle\sum_{\nu=a+1}^{a+n} x_\nu \|^2 \le \sigma^2 n \tag{3.2}$$

for all $a \ge 0$, $n \ge 1$. *Then with probability* 1

$$\limsup_{N\to\infty} \| \textstyle\sum_{\nu \le N} x_\nu \| / a_N \le 2000\sigma/\delta.$$

The proof of Theorem 5 is divided into three sections.

3.1. *Bounds on the* $(2 + \delta)$th *moments of partial sums.*

LEMMA 3.1. *Let $\{\eta_n, n \ge 1\}$ be a sequence of B-space valued random variables, satisfying the hypotheses of Theorem* 5. *We assume that* $0 < \delta \le 1$. *Suppose that* (3.1) *holds. Then*

$$E \| \textstyle\sum_{\nu=a+1}^{a+n} \eta_\nu \|^{2+\delta} \ll n^{1+\delta/2}(\sigma^{2+\delta} + b).$$

where the constant implied by $\ll$ only depends on ϵ, δ and the constant implied by $\ll$ in (3.1). *Here b denotes the uniform bound on the* $(2 + \delta)$th *moments.*

This lemma is a variant of Ibragimov's (1962) Lemma 1.9. Its proof is a minor modification of the proof of Theorem 3.1 of Serfling (1968). Without loss of generality we assume $\sigma^{2+\delta} + b \le 1$. Since $\delta \le 1$ and since his $\beta = \tfrac{1}{2}\delta$ we can replace his ϵ by δ. For given n we define

$$k = [n^{1/2-\epsilon/8}], \qquad m = [\tfrac{1}{2}n] - k$$

so that

$$\phi(2k) \ll n^{-(1+2/\delta)}. \tag{3.3}$$

Here and throughout this section ϵ is the same as in (3.1). Define

$$R_a = \textstyle\sum_{\nu=a+1}^{a+m} \eta_\nu, \qquad S_a = \sum_{\nu=a+m+2k+1}^{a+2m+2k} \eta_\nu.$$

Then each of these two sums has m terms and thus by (3.2) we have uniformly in $a \ge 0$

$$E \| S_a \|^2 \le m.$$

Moreover, we observe that by (3.3), by Serfling's (1968) Theorem 2.2 with $\alpha = p = 1 + \frac{1}{2}\delta$, $q = 1 + 2/\delta$ and by Minkowski's inequality

$$\Delta = E(\| S_a \|^2 \mid \mathcal{M}_1^{a+m}) - E \| S_a \|^2$$

satisfies

$$E \mid \Delta \mid^{1+\delta/2} \leq 4\phi(2k)^{\delta/2} E \| S_a \|^{2+\delta} \ll n^{1+\delta/2}.$$

Thus by Serfling's argument

$$E \| R_a + S_a \|^{2+\delta} \leq m^{1+\delta/2} a_m (2^{1/2} - z_0^{-1})^{2+\delta}.$$

Consequently, using Minkowski's inequality (here and in the next five lines a_m has the same meaning as in Serfling's paper)

$$\begin{aligned} \| \textstyle\sum_{\nu=a+1}^{a+n} \eta_\nu \|_{2+\delta} &\leq \| R_a + S_a \|_{2+\delta} + \| \textstyle\sum_{\nu=a+m+1}^{a+m+2k} \eta_\nu \|_{2+\delta} + \| \eta_{a+n} \|_{2+\delta} \\ &\leq m^{1/2} a_m^{1/(2+\delta)} (2^{1/2} - z_0^{-1}) + 2k + 1 \\ &\leq m^{1/2} a_m^{1/(2+\delta)} (2^{1/2} - z_0^{-1} + 3m^{-\epsilon/8}) \end{aligned}$$

since $a_m \geq 1$. Thus for all $n \geq N_0 = (6z_0)^{8/\epsilon}$

$$E \| \textstyle\sum_{\nu=a+1}^{a+n} \eta_\nu \|^{2+\delta} \leq n^{1+\delta/2} a_m.$$

The remainder of the proof is the same as in Serfling (1968).

We note that the idea of separating the blocks R_a and S_a in this context goes back to Doob (1953), pages 225–227.

3.2. *An exponential bound.* The proof of Theorem 5 is based on the following proposition.

PROPOSITION 3.1. *Let $R \geq 1$ and assume the hypotheses of Theorem 5. Suppose that the bound b for the $(2 + \delta)$th moments is not less than 1. Then there exists a constant $C \geq 1$ depending only on ϵ, δ and the constant implied by $\ll$ in (3.1) such that for all $n \geq 0$ and $a \geq 0$*

$$\begin{aligned} &P\{\| \textstyle\sum_{\nu=a+1}^{a+2^n} x_\nu \| \geq 10R\sigma(2^n \log n)^{1/2}\} \\ &\qquad \leq C(2^{-2n} + R^{-2} 2^{-n\delta/20}(1 + b\sigma^{-2-\delta}) + \exp(-1.1\, R^{\delta/(2+\delta)} \log n)). \end{aligned}$$

Without loss of generality we can assume $a = 0$ as well as

$$R^{-2}(1 + b\sigma^{-2-\delta}) < 2^{n\delta/20} \tag{3.4}$$

since otherwise there is nothing to prove. We divide the proof of Proposition 3.1 into several steps and formulate them as lemmas. Let

$$s = \left[\frac{4}{5} n\right] \quad \text{and} \quad t = \left[\frac{2}{5} n\right]. \tag{3.5}$$

We introduce long blocks $H_1, \cdots, H_k$ of consecutive integers, each having length $2^s - 2^t$ and short blocks $I_1, \cdots, I_k$ of length 2^t each, leaving no gaps between the blocks. The order is $H_1, I_1, \cdots, H_k, I_k$. We define k by

$$k = 2^{n-s} (\ll 2^{n/5}) \tag{3.6}$$

so that

$$\textstyle\bigcup_{j \leq k} H_j \cup I_j = [1, 2^n] \cap \mathbb{Z}.$$

We write

$$y_j = \textstyle\sum_{\nu \in H_j} x_\nu \qquad \text{and} \qquad z_j = \textstyle\sum_{\nu \in I_j} x_\nu \qquad\qquad 1 \leq j \leq k$$

and note that by Markov's and Minkowski's inequalities and by (3.2), (3.5) and (3.6)

$$\begin{aligned} P\{\| \textstyle\sum_{j \leq k} z_j \| \geq \frac{1}{2} R\sigma 2^{n/2}\} &\leq (\tfrac{1}{2} R\sigma 2^{n/2})^{-2} (\textstyle\sum_{j \leq k} (E \| z_j \|^2)^{1/2})^2 \\ &\ll R^{-2} 2^{-n/5}. \end{aligned} \tag{3.7}$$

We also note that

$$\sum_{j\le k}(y_j + z_j) = \sum_{\nu\le 2^n} x_\nu$$

and that thus by the triangle and Hölder's inequalities and by (3.2), (3.12) and (3.6)

$$\begin{aligned} E\,\|\textstyle\sum_{j\le k} y_j\| &\le \textstyle\sum_{j\le k} E\,\|z_j\| + \sigma 2^{n/2} \\ &\le \sigma k\cdot 2^{t/2} + \sigma 2^{n/2} \ll \sigma 2^{n/2} \end{aligned} \tag{3.8}$$

By Theorem 2 of Berkes and Philipp (1979) there exist independent random variables Y_j having the same distribution as y_j such that

$$P\{\|Y_j - y_j\| \ge 6\phi(2^t)\} \le 6\phi(2^t). \tag{3.9}$$

Thus by (3.1), (3.4), (3.5) and (3.6)

$$\begin{aligned} P\{\|\textstyle\sum_{j\le k}(Y_j - y_j)\| \ge \tfrac12 R^{2/(2+\delta)}\sigma 2^{n/2}\} &\le \textstyle\sum_{j\le k} P\{\|Y_j - y_j\| \ge \tfrac12 R^{2/(2+\delta)}\sigma 2^{n/2}k^{-1}\} \\ &\ll k\cdot\phi(2^t) \ll 2^{-2n}. \end{aligned} \tag{3.10}$$

We now truncate the random variables Y_j by defining

$$\begin{aligned} w_j &= Y_j \qquad \text{if } \|Y_j\| < R^{2/(2+\delta)}\sigma 2^{n/2}(\log n)^{-1/2} \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{3.11}$$

Since Y_j and y_j have the same distribution we obtain from Lemma 3.1, (3.4) and (3.5)

$$\begin{aligned} P\{Y_j \ne w_j\} &= P\{\|y_j\| \ge R^{2/(2+\delta)}\sigma 2^{n/2}(\log n)^{-1/2}\} \\ &\le R^{-2}\sigma^{-2-\delta}2^{-n(1+\delta/2)}(\log n)^{1+\delta/2}E\,\|y_j\|^{2+\delta} \\ &\ll R^{-2}2^{-n(1+\delta/2)/5}(1 + b\sigma^{-2-\delta}) \ll 2^{-n(1+\delta/4)/5}. \end{aligned} \tag{3.12}$$

LEMMA 3.2 $E\,\|\sum_{j\le k} w_j\| \ll R^{2/(2+\delta)}\sigma 2^{n/2}$

PROOF. Let

$$A = \{\|\textstyle\sum_{j\le k}(y_j - w_j)\| \ge R^{2/(2+\delta)}\sigma 2^{n/2}\}$$

$$B = \{\|\textstyle\sum_{j\le k}(y_j - Y_j)\| \ge R^{2/(2+\delta)}\sigma 2^{n/2}\}$$

and

$$C = \bigcup_{1\le\nu\le k}\bigcup_{1\le j_1<\cdots j_\nu\le k}(C(j_1,\cdots,j_\nu)\cap B^c)$$

where

$$C(j_1,\cdots,j_\nu) = \{w_{j_1}\ne Y_{j_1},\cdots, w_{j_\nu}\ne Y_{j_\nu}, w_j = Y_j \quad \text{for all } j\ne j_i\ (1\le i\le\nu)\}.$$

Obviously $A\subset B\cup C$. By (3.8)

$$\int_{A^c}\|\textstyle\sum_{j\le k} w_j\| \le \int_{A^c}\|\textstyle\sum_{j\le k} y_j\| + R^{2/(2+\delta)}\sigma 2^{n/2} \ll R^{2/(2+\delta)}\sigma 2^{n/2} \tag{3.13}$$

and by (3.6), (3.10) and (3.11)

$$\int_B \|\textstyle\sum_{j\le k} w_j\| \le k\cdot 2^{n/2}R^{2/(2+\delta)}P(B)\sigma \ll 2^{n/5+n/2}R^{2/(2+\delta)}2^{-2n}\sigma \ll R^{2/(2+\delta)}\sigma. \tag{3.14}$$

By (3.12) and independence we have for each ν-tuple $j_1 < \cdots < j_\nu$

$$P\{C(j_1,\cdots,j_\nu)\} \ll 2^{-\nu n(1+\delta/4)/5} \tag{3.15}$$

and on $B^c \cap C(j_1, \cdots, j_\nu)$ we have

$$\begin{aligned} \| \textstyle\sum_{j\le k} w_j \| &\le \| \textstyle\sum_{j\le k} Y_j \| + \textstyle\sum_{i\le\nu} \| Y_{j_i} \| \\ &\le \| \textstyle\sum_{j\le k} y_j \| + R^{2/(2+\delta)}\sigma 2^{n/2} + \textstyle\sum_{i\le\nu} \| Y_{j_i} \|. \end{aligned} \tag{3.16}$$

By independence, Cauchy's inequality, (3.15), (3.12), (3.5), (3.2) and since Y_j and y_j have the same distribution we have

$$\int_{C(j_1,\cdots,j_\nu)} \| Y_{j_i} \|$$

$$= P\{w_{j_l} \ne Y_{j_l},\ 1 \le l \le \nu,\ l \ne i;\ w_j = Y_j \quad \text{for all} \quad j \ne j_l,\ 1 \le l \le \nu\} \int_{\{w_{j_i} \ne Y_{j_i}\}} \| Y_{j_i} \|$$

$$\ll 2^{-(\nu-1)n(1+\delta/4)/5} \cdot P^{1/2}\{w_{j_i} \ne Y_{j_i}\} \cdot E^{1/2} \| Y_{j_i} \|^2$$

$$\ll 2^{-(\nu-1)n(1+\delta/4)/5} \cdot 2^{-n(1+\delta/4)/10} 2^{2n/5}\sigma.$$

Consequently by (3.8), (3.6) and (3.16)

$$\begin{aligned} \int_C \| \textstyle\sum_{j\le k} w_j \| &\le \int_C \| \textstyle\sum_{j\le k} y_j \| + R^{2(2+\delta)}\sigma 2^{n/2} \\ &\quad + \textstyle\sum_{\nu\le k} k^\nu \nu 2^{-(\nu-1)n(1+\delta/4)/5} \cdot 2^{-n(1+\delta/4)/10} 2^{2n/5}\sigma \\ &\ll R^{2/(2+\delta)}\sigma 2^{n/2} + \sigma 2^{n/2} \cdot 2^{n\delta/40} \textstyle\sum_{1\le\nu\le k} \nu 2^{-n\nu\delta/25} \ll R^{2/(2+\delta)}\sigma 2^{n/2}. \end{aligned}$$

The lemma follows now from (3.13), (3.14) and the last estimate.

The following lemma is essentially due to Kuelbs (1977).

LEMMA 3.3 *Let $\{w_j, j \le k\}$ be a sequence of independent B-valued random variables with*

$$\| w_j \| \le cb_k \qquad 1 \le j \le k.$$

Let

$$T_k = \textstyle\sum_{j\le k} w_j.$$

Then

$$\begin{aligned} P\{\| T_k \| \ge 2\epsilon b_k\} &\le \exp\{-\epsilon^2 + \tfrac12 \epsilon^2(1 + \tfrac12 \epsilon c)b_k^{-2} \textstyle\sum_{j\le k} E \| w_j \|^2 \\ &\quad + \tfrac12 \epsilon b_k^{-1} E \| T_k \|\} && \text{if } \epsilon c \le 1 \\ &\le \exp\{-\epsilon c^{-1}(1 - \tfrac34 b_k^{-2} \textstyle\sum_{j\le k} E \| w_j \|^2) \\ &\quad + \tfrac12 c^{-1} b_k^{-1} E \| T_k \|\} && \text{if } \epsilon c > 1. \end{aligned}$$

PROOF. By relation (2.4) of Kuelbs (1977) we have for all t with $tc \le 1$

$$\begin{aligned} P\{\| T_k \| > 2\epsilon b_k\} &\le \exp(-\epsilon t)E\{\exp(\tfrac12 b_k^{-1} t \| T_k \|)\} \\ &\le \exp\{-\epsilon t + \tfrac12 t^2(1 + \tfrac12 tc)b_k^{-2} \textstyle\sum_{j\le k} E \| w_j \|^2 + \tfrac12 t b_k^{-1} E \| T_k \|\}. \end{aligned}$$

The lemma follows if we set $t = \epsilon$ or $t = 1/c$ according as $\epsilon c \le 1$ or $\epsilon c > 1$.

We now can finish the proof of Proposition 3.1. We first observe that by (3.11), (3.2) and the definition of y_j

$$E \| w_j \|^2 \le E \| Y_j \|^2 = E \| y_j \|^2 \le 2^s\sigma^2.$$

Thus by (3.5)

$$\textstyle\sum_{j\le k} E \| w_j \|^2 \le k \cdot 2^s\sigma^2 \le 2^n\sigma^2. \tag{3.17}$$

We now apply Lemma 3.3 with $b_k = \sigma 2^{n/2}$, $c = R^{2/(2+\delta)}(\log n)^{-1/2}$ and $\epsilon = \frac{9}{2}R(\log n)^{1/2}$. Then $\epsilon c > 1$ and thus by (3.17) and Lemma 3.2

$$P\{\| \sum_{j\le k} w_j \| \ge 9R\sigma(2^n \log n)^{1/2}\} \le \exp\{- \tfrac{9}{2} R^{\delta/(2+\delta)} \log n \cdot (1 - \tfrac{3}{4}) + O(\log^{1/2} n)\}$$
$$\ll \exp(-1.1\, R^{\delta/(2+\delta)} \log n).$$

Here the constant implied by O only depends on the constant implied by $\ll$ in Lemma 3.2. Thus by (3.10), (3.12), (3.6) and (3.7)

$$\begin{aligned} P\{\| \textstyle\sum_{\nu\le 2^n} x_\nu \| \ge 10R\sigma(2^n \log n)^{1/2}\} & \\ \le P\{\| \textstyle\sum_{j\le k}(Y_j - y_j)\| \ge \tfrac{1}{2} R\sigma 2^{n/2}\} &+ P\{\| \textstyle\sum_{j\le k} z_j \| \ge \tfrac{1}{2} R\sigma 2^{n/2}\} \\ + \textstyle\sum_{j\le k} P\{w_j \ne Y_j\} &+ P\{\| \textstyle\sum_{j\le k} w_j \| \ge 9R\sigma(2^n \log n)^{1/2}\} \\ \ll 2^{-2n} + R^{-2} 2^{-n/5} &+ R^{-2}(1 + b\sigma^{-2-\delta}) 2^{-n\delta/20} + \exp(- 1.1 R^{\delta/(2+\delta)} \log n). \end{aligned}$$

The proposition follows now since $0 < \delta \le 1$.

3.3. Proof of Theorem 5. We can assume without loss of generality $\sigma = 1$. We write

$$F(M, N) = \| \textstyle\sum_{\nu=M+1}^{M+N} x_\nu \|. \tag{3.18}$$

Let $N \ge 1$ be given and let n be the largest integer with $2^n \le N$. Put $p = [3n/8] + 1$. Then we can write

$$N = 2^n + \textstyle\sum_{l=p}^{n} \epsilon_l 2^{l-1} + \theta 2^p$$

where $\epsilon_l = 0$ or 1 and $|\theta| \le \frac{1}{2}$. Thus there exist integers m_l with $0 \le m_l < 2^{n-l}$ $(p \le l \le n)$ such that (for the details see S. and L. Gaal (1964), page 139)

$$F(0, N) \le F(0, 2^n) + \textstyle\sum_{l=p}^{n} F(2^n + m_l 2^l, 2^{l-1}) + \sum_{\nu=2^n+m_p 2^p+1}^{2^n+m_p 2^p+2^p} \| x_\nu \|. \tag{3.19}$$

We now define the events

$$E_n = \{F(0, 2^n) \ge 10(2^n \log n)^{1/2}\}$$
$$G_n(m, l) = \{F(2^n + m2^l, 2^{l-1}) \ge 10\cdot 2^{n/2+\delta(l-n)/80} \log^{1/2} n\}$$
$$G_n = \textstyle\bigcup_{l=p}^{n} \bigcup_{m\le 2^{n-l}} G_n(m, l)$$
$$H_n(m) = \{\textstyle\sum_{\nu=2^n+m2p+1}^{2^n+m2^p+2^p} \| x_\nu \| \ge 2^{n/2}\}$$

and

$$H_n = \textstyle\bigcup_{m\le 2^{n-p}} H_n(m).$$

Lemma 3.4. *With probability* 1 *only finitely many of the events* E_n, G_n *and* H_n *occur.*

Proof. We put $R = 1$ in Proposition 3.1 and obtain

$$P(E_n) \ll 2^{-n\delta/20} + \exp(- 1.1 \log n) \ll n^{-1.1}. \tag{3.20}$$

Similarly if we put $R = 2^{(n-l)(1/2-\delta/80)}$ we obtain

$$P(G_n(m, l)) \ll 2^{-(n-l)(1-\delta/40)-l\delta/20} + 2^{-2l} + \exp(- 1.1\, 2^{\delta(n-l)/8} \log l).$$

Thus

$$\begin{aligned} P(G_n) &\ll \textstyle\sum_{3n/8\le l\le n} 2^{(n-l)\delta/40-l\delta/20} + \sum_{3n/8\le l\le n} 2^{-2l}\cdot 2^{n-l} \\ &\quad + \textstyle\sum_{3n/8\le l\le n} \exp(- 1.1\cdot 2^{\delta(n-l)/8} \log l + n - l) \\ &\ll n^{-1.1}. \end{aligned} \tag{3.21}$$

Finally,

$$P(H_n(m)) \leq P\{|\sum (\| x_\nu \| - E \| x_\nu \|)| \geq \tfrac{1}{2} 2^{n/2}\} + P\{\sum E \| x_\nu \| \geq \tfrac{1}{2} 2^{n/2}\}$$

where both sums are extended over appropriate intervals of length 2^p. The second probability is zero since $p \leq \frac{3}{8}n + 1$. The first probability is by Lemma 3.1.

$$\ll 2^{-n(1+\delta/2)} \cdot 2^{p(1+\delta/2)}.$$

Thus

$$P(H_n) \ll 2^{-(n-p)\delta/2} \ll 2^{-n\delta/4}. \tag{3.22}$$

The lemma follows now from (3.20)–(3.22) and the Borel-Cantelli lemma.

To finish the proof of Theorem 5 we observe that by (3.19) and by Lemma 3.4 we have with probability 1 for all $N \geq N_0(\omega)$

$$F(0, N) \leq 10(2^n \log n)^{1/2} + \textstyle\sum_{3n/8 \leq l \leq n} 10 \cdot 2^{n/2+\delta(l-n)/80} \log^{1/2} n + 2^{n/2}$$

$$\leq 2000/\delta \cdot (2^n \log n)^{1/2} \leq 2000/\delta \cdot (N \log \log N)^{1/2}.$$

We divide by $(N \log \log N)^{1/2}$, take the $\limsup_{N\to\infty}$ and obtain the result.

4. Proof of Theorems 1 and 2. The goal of this section is a little ambitious. We first shall prove the rather general Theorem 6 below from which Theorem 1 and half of Theorem 3 will follow fairly quickly. Moreover, we have arranged the proof of Theorem 6 so that first the proof will require only minor modifications to yield Theorem 2 and also that Theorems 1 and 2 can be easily extended to random variables satisfying a strong mixing condition once Proposition 3.1 has been generalized accordingly. We hope that the exposition is sufficiently translucent so that the reader will not judge our goal as over-ambitious.

Before we start we would like to dispose of an idea for the proof of Theorem 1 which does not work. Considering conditions (1.7)–(1.9) and the statements of Theorems 4 and 5, the general direction of how to proceed from there on appears to be quite obvious. For given $\rho > 0$ the random variables $x_j - \Lambda_\rho(x_j)$ satisfy the hypotheses of Theorem 5 and thus

$$\limsup_{n\to\infty} \| \textstyle\sum_{j\leq n} x_j - \Lambda_\rho(x_j) \| / a_n \ll \rho^{1/2} \qquad \text{a.s.}$$

Since the random variables $\Lambda_\rho(x_j)$ are finite dimensional and mixing Theorem 4 gives a Brownian motion $\{X_\rho(t), t \geq 0\}$ such that

$$\| \textstyle\sum_{j\leq t} \Lambda_\rho(x_j) - X_\rho(t) \| \ll t^{1/2-\lambda} \qquad \text{a.s.}$$

Hence

$$\limsup_{t\to\infty} \| \textstyle\sum_{j\leq t} x_j - x_\rho(t) \| (t \log \log t)^{-1/2} \ll \rho^{1/2} \qquad \text{a.s.} \tag{4.1}$$

Since $\rho > 0$ is arbitrary it follows from the remarks after Theorem 3 that (4.1) still implies the compact as well as the functional law of the iterated logarithm for $\{x_j, j \geq 1\}$. However, (4.1) does not guarantee the existence of a universal Brownian motion $\{X(t), t \geq 0\}$ satisfying (1.11). As a matter of fact this is just one thing the example mentioned in Section 1 and carried out in Section 5, is designed to demonstrate. Hence a substantially different idea has to be introduced. It consists, in essence, of a utilization of the maps Π_N described in Lemma 2.1 of Kuelbs (1976a).

We now formulate the above-mentioned result which also serves as an outline of the proofs of our main theorems.

THEOREM 6. *Let $\{x_j, j \geq 1\}$ be a weak sense stationary sequence of B-valued random variables centered at expectations and with $(2 + \delta)$th moments with $0 < \delta \leq 1$ uniformly bounded by 1. Suppose that $\{x_j, j \geq 1\}$ satisfies a strong mixing condition (2.1) with rate of decay given by (2.2). Moreover, let α be given by Lemma 2.5, i.e., $\alpha = \varepsilon\delta/8$ and recall (2.6). Suppose that there exist constants A and C such that for all $R \geq 1$, $a \geq 0$ and $n \geq 0$*

$$P\{\| \textstyle\sum_{\nu=a+1}^{a+2^n} x_\nu \| \geq AR(2^n \log n)^{1/2}\} \leq C(2^{-2n} + R^{-2} 2^{-n\alpha/20} + \exp(-1.1 R^{\alpha/(2+\alpha)} \log n)). \tag{4.2}$$

Next, let μ be any mean zero Gaussian measure on B and let Π_N be the maps obtained from μ as defined in Lemma 2.1 *of Kuelbs* (1976a). *Then $\{\Pi_N x_j, j \geq 1\}$ is a weak sense stationary sequence of random variables centered at expectations with* $(2 + \delta)$th *moments uniformly bounded by $\| \Pi_N \|^{2+\delta}$. Hence by Proposition* 2.1 *for any fixed $a \geq 0$ the sequence $\{\Pi_N x_{j+a}, j \geq 1\}$ satisfies the central limit theorem with limiting Gaussian measure λ_N (say), independent of a. We suppose that μ satisfies $\mu^{\Pi_N} = \lambda_N$ for $N \geq N_0$.*

Finally, let K be the unit ball of the Hilbert space H_μ as defined in Lemma 2.1 *of Kuelbs* (1976a). *Suppose that with probability* 1

$$\lim_{n\to\infty} \| S_n/a_n - K \| = 0. \tag{4.3}$$

Then the conclusion of Theorem 1 *holds.*

The proof of Theorem 6 is carried out in the next three subsections. In Section 4.4 we deduce Theorem 1 from Theorem 6 and in Section 4.5 we prove Theorem 2 by a minor modification of the argument.

4.1. *Approximation by finite-dimensional random variables.* Let Π_N be the maps obtained from μ as described in Lemma 2.1 of Kuelbs (1976a). Then by relation (2.4) of Kuelbs (1976a)

$$\dim \Pi_N B = \min(N, \dim H_\mu). \tag{4.4}$$

LEMMA 4.1. *Given $\eta > 0$ there is an N such that*

$$\limsup_{n\to\infty} a_n^{-1} \| \textstyle\sum_{\nu\leq n} (x_\nu - \Pi_N x_\nu) \| < \eta \qquad \text{a.s.}$$

PROOF. Let I be the identity map on B. Since the map $I - \Pi_N$ is continuous

$$\| a_n^{-1}(S_n - \Pi_N S_n) - (I - \Pi_N)K \| \to 0 \qquad \text{a.s.} \tag{4.5}$$

Now by relation (3.7) of Kuelbs (1976a)

$$(I - \Pi_N)K \subseteq \{x \in B: \| x \| < \eta\}$$

for all sufficiently large N. This together with (4.5) implies the lemma.

LEMMA 4.2. *Let $X(t)$ be Brownian motion on B having the same covariance structure as μ. Then given $\eta > 0$ there is an N such that*

$$\limsup_{t\to\infty} (t \log\log t)^{-1/2} \| X(t) - \Pi_N X(t) \| < \eta \qquad \text{a.s.}$$

PROOF. We apply Theorem 4.1 of Kuelbs (1977) to the increments of $X(n)$ and obtain

$$\lim_{n\to\infty} \| X(n)/a_n - K \| = 0 \qquad \text{a.s.}$$

Let

$$Z_n = \max_{n\leq t\leq n+1} \| X(t) - X(n) \| \qquad n \geq 0.$$

Then $\{Z_n, n \geq 0\}$ is a sequence of independent identically distributed random variables. By Lévy's maximal inequality for B-valued Brownian motion we have

$$P\{Z_n \geq \lambda\} = P\{Z_0 \geq \lambda\} \leq 2P\{\| X(1) \| \geq \lambda\}.$$

Applying the Fernique-Landau-Shepp theorem (see Fernique (1970)) to the distribution of $\| X(1) \|$ we have by a standard application of the Borel-Cantelli lemma

$$\max_{n\leq t\leq n+1} \| X(t) - X(n) \| \ll n^{1/4} \qquad \text{a.s.} \tag{4.6}$$

Thus

$$\lim_{t\to\infty} \| (2t \log\log t)^{-1/2} X(t) - K \| = 0 \qquad \text{a.s.}$$

The remainder of the proof of the lemma follows verbatim the proof of Lemma 4.1 except that we replace S_n by $X(t)$ and a_n by a_t.

4.2. *Introduction of the blocks.* Let α be given by Lemma 2.5, i.e., $\alpha = \varepsilon\delta/8$ and let N_k be the largest integral power of 2 not exceeding $k^{-1+\alpha/10}\exp(k^{\alpha/10})$. As in Section 2.3 we define inductively blocks of consecutive integers, H_k and I_k, leaving no gaps between the blocks, by setting

$$\text{card } H_k = N_k, \text{ card } I_k = [\exp(\tfrac{1}{4}k^{\alpha/10})] \qquad k \geq 1. \tag{4.7}$$

Let

$$t_k = \textstyle\sum_{j\leq k} \text{card}(H_j \cup I_j). \tag{4.8}$$

Then

$$\exp(k^{\alpha/10}) \ll t_k \ll \exp(k^{\alpha/10}). \tag{4.9}$$

Of course, the blocks H_k have nothing to do with the Hilbert spaces H_μ or $H_{\mathscr{L}(x_1)}$. If $\{x_j, j \geq 1\}$ is a sequence of ϕ-mixing random variables with rate of decay given by (1.6) we shall replace α by δ throughout this and the next two subsections. This will be used in the proof of Theorem 2.

LEMMA 4.3. *As* $k \to \infty$

$$\| \textstyle\sum_{\nu\in I_k} x_\nu \| \ll t_k^{1/3} \qquad \text{a.s.}$$

PROOF. Since the $(2 + \delta)$th moments of x_ν are uniformly bounded by 1, so are the first moments. Hence by (4.7), (4.9) and Markov's inequality

$$P\{\| \textstyle\sum_{\nu\in I_k} x_\nu \| \geq t_k^{1/3}\} \ll \exp(-\tfrac{1}{3}k^{\alpha/10}) \textstyle\sum_{\nu\in I_k} E\| x_\nu \| \ll k^{-2}.$$

PROPOSITION 4.1. *We have with probability* 1 *as* $k \to \infty$

$$\max_{t_k<t\leq t_{k+1}} \| \textstyle\sum_{t_k<\nu\leq t} x_\nu \| \ll (t_k/\log t_k)^{1/2}.$$

For the proof we combine the arguments used in the proofs of Proposition 2.2 and Section 3.3. We define $F(r, s)$ by (2.25), replacing η_ν by x_ν. Put $N = [t]$ and define $n(N)$ by (2.26). Let $n_k = n(t_{k+1})$ and $p = p_k = [3n_k/8] + 1$. Then as in (2.27) and (3.19) we have for each N with $t_k < N \leq t_{k+1}$

$$F(0, N - t_k) \leq \textstyle\sum_{p\leq l\leq n} F(m_l 2^{l+1}, 2^l) + \sum_{\nu=t_k+m_p' 2^p+1}^{t_k+(m_p'+1)2^p} \| x_\nu \| \tag{4.10}$$

where $0 \leq m_l < 2^{n-l}$ $(p \leq l \leq n)$ and $0 \leq m_p' \leq 2^{n_k-p+1}$. Next we define the events

$$G_k(m, l) = \{F(m2^{l+1}, 2^l) \geq A(t_k/\log^3 t_k)^{1/2}\}$$

$$G_k = \textstyle\bigcup_{p\leq l\leq n_k} \bigcup_{m\leq 2^{n_k-l}} G_k(m, l)$$

$$H_k(m) = \{\textstyle\sum_{\nu=t_k+m2^p+1}^{t_k+(m+1)2^p} \| x_\nu \| \geq (t_k/\log t_k)^{1/2}\}$$

$$H_k = \textstyle\bigcup_{m\leq 2^{n_k-p+1}} H_k(m).$$

LEMMA 4.4. *With probability* 1 *only finitely many of the events* G_k *and* H_k *occur.*

PROOF. We apply (4.2) with $R = t_k^{1/2}\, 2^{-l/2}(\log t_k)^{-2}$ and obtain

$$\begin{aligned} P(G_k(m, l)) &\ll 2^{-2l} + t_k^{-1}2^l \log^4 t_k \cdot 2^{-l\alpha/20} \\ &\quad + \exp(-1.1(t_k^{1/2}2^{-l/2} (\log t_k)^{-2})^{\alpha/(2+\alpha)} \log l). \end{aligned} \tag{4.11}$$

Consequently, by a straight-forward calculation using (4.7) and (4.9)

$$P(G_k) \ll \exp(-\tfrac{1}{10}k^{\alpha/10}) + \exp\left(-\frac{\alpha}{100}k^{\alpha/10}\right) + \exp(-k^{\alpha/10}) \tag{4.12}$$

where each term in (4.12) represents a bound for the sum of the corresponding terms in (4.11).
Finally,

$$P(H_k(m)) \leq P\{|\Sigma(\|x_\nu\| - E\|x_\nu\|)| \geq \tfrac{1}{2}(t_k/\log t_k)^{1/2}\} + P\{\Sigma E\|x_\nu\| \geq \tfrac{1}{2}(t_k/\log t_k)^{1/2}\}$$

where both sums are extended over appropriate intervals of length 2^p. The second probability is zero since $p \leq 3n_k/8 + 1$. By Lemma 2.5 the first probability is bounded by

$$\ll (t_k/\log t_k)^{-1-(1/2)\alpha} \cdot 2^{p(1+(1/2)\alpha)}.$$

(If $\{x_j, j \geq 1\}$ is ϕ-mixing we apply Lemma 3.1 instead.) Hence by (4.7) and (4.9)

$$P(H_k) \ll \exp(-\tfrac{1}{16}\alpha k^{\alpha/10}).$$

The lemma follows now from this last estimate, (4.12) and the Borel-Cantelli lemma. □

We now can finish the proof of Proposition 4.1. By (4.10) we have with probability 1

$$\max_{t_k < N \leq t_{k+1}} F(0, N - t_k) \ll n_k(t_k/\log^3 t_k)^{1/2} \ll (t_k/\log t_k)^{1/2}.$$

LEMMA 4.5. *We have with probability* 1 *as* $k \to \infty$

$$\max_{t_k \leq t \leq t_{k+1}} \|X(t) - X(t_k)\| \ll (t_k/\log t_k)^{1/2}.$$

This follows from the Fernique-Landau-Shepp theorem (see Fernique (1970)) by standard calculations. But it also follows from (4.6) and from Proposition 4.1 applied to the sequence of increments $X(n) - X(n-1)$.

4.3. *Conclusion of the proof of Theorem* 6. Let $d_k \uparrow \infty$ subject to the conditions

$$d_k \leq k^{\alpha(1-\beta)/10}, \quad \|\Pi_{d_k}\|_1 \ll \exp(k^{\alpha(1-\beta)/10}). \tag{4.13}$$

Here we restrict β to $0 < \beta \leq \delta/20$. Recall that in case of ϕ-mixing random variables α is to be replaced by δ.

We put $d = \min(d_k, \dim H_\mu)$. Then $\Pi_{d_k}B = \Pi_{d_k}H_\mu$ can be viewed as a Euclidean space $\mathbb{R}^d$ with metric induced by the norm in H_μ. As usual, we denote the metric $\|\cdot\|_{H_\mu}$ by $|\cdot|$. Throughout the remainder of this section we will frequently identify $\Pi_{d_k}B$ as $\mathbb{R}^d$ without notice. We define

$$\xi_\nu = \Pi_{d_k}x_\nu \qquad \text{if} \quad \nu \in H_k \cup I_k. \tag{4.14}$$

Then $\{\xi_\nu, \nu \in H_k\}$ is a weak sense stationary sequence of random variables centered at expectations with $E|\xi_\nu|^{2+\delta} \ll \exp(k^{\alpha(1-\beta/2)/10})$ and satisfying a strong mixing condition with the same rate of decay given by (2.2). We also define

$$X_k = N_k^{-1/2} \textstyle\sum_{\nu \in H_k} \xi_\nu. \tag{4.15}$$

We now apply Theorem 1 of Berkes and Philipp (1979) to the sequence $\{X_k, k \geq 1\}$ and to $\{\mathscr{F}_k, k \geq 1\}$ where $\mathscr{F}_k = \mathscr{M}_1^{t_{k-1}+N_k}$. The random vectors X_k have dimension not exceeding d_k. We put

$$T_k = k^3. \tag{4.16}$$

By Lemma 2.2, (4.7), (4.8) and (2.2)

$$\begin{aligned} E|E\{\exp(i\langle u, X_k\rangle)|\mathscr{F}_{k-1}\} - E\{\exp(i\langle u, X_k\rangle)\}| &\ll \rho(\exp(\tfrac{1}{4}k^{\alpha/10})) \\ &\ll \exp(-\tfrac{1}{4}k^{\alpha/10}). \end{aligned} \tag{4.17}$$

By Proposition 2.1, (4.6) and (4.7) we have for all $|u| \leq T_k \ll N_k^{\alpha/64}$

$$E\{\exp(i\langle u, X_k\rangle)\} - \exp\{-\tfrac{1}{2}\langle u, \Gamma_d u\rangle\} \ll N_k^{-\alpha/32}b \ll \exp(-10^{-2}\alpha k^{\alpha/10}). \tag{4.18}$$

Here Γ_d is the $d \times d$ identity matrix since, by hypothesis, Γ_d is the covariance matrix of the Gaussian measure μ^{Π_d}. Thus we can choose

$$\lambda_k = \text{const} \exp(-10^{-2}\alpha k^{\alpha/10}). \tag{4.19}$$

As in Section 3 of the paper by Berkes and Philipp (1979) one can show that

$$G_k\{|u| \geq \tfrac{1}{4}T_k\} \ll e^{-k} = \delta_k, \qquad \text{say}, \tag{4.20}$$

where G_k is a multivariate normal distribution with mean zero and covariance matrix Γ_d. Hence by Theorem 1 of Berkes and Philipp (1979) we can redefine the sequence $\{X_k, k \geq 1\}$ without changing its distribution on a richer probability space on which there exists a sequence $\{Y_k, k \geq 1\}$ of independent random vectors with distribution G_k, having values in $\mathbb{R}^d$ and satsifying

$$P\{|X_k - Y_k| \geq \alpha_k\} \leq \alpha_k \tag{4.21}$$

with

$$\alpha_k \ll d_k T_k^{-1} \log T_k + \lambda_k^{1/2} T_k^{d_k} + \delta_k \ll k^{-2} \tag{4.22}$$

by (4.13), (4.16), (4.19) and (4.20). Hence with probability 1

$$|X_k - Y_k| \ll k^{-2}$$

or

$$|\textstyle\sum_{\nu \in H_k} \xi_\nu - N_k^{1/2} Y_k| \ll N_k^{1/2} k^{-2} \ll t_k^{1/2} k^{-2}. \tag{4.23}$$

Further by (4.20)

$$|Y_k| \ll T_k \qquad \text{a.s.}$$

Consequently by (4.23), (4.7) and (4.8) we have with probability 1

$$\begin{aligned} &|\textstyle\sum_{\nu \in H_k} \xi_\nu - (t_k - t_{k-1})^{1/2} Y_k| \\ &\quad \leq |\textstyle\sum_{\nu \in H_k} \xi_\nu - N_k^{1/2} Y_k| + |Y_k| \, |(t_k - t_{k-1})^{1/2} - N_k^{1/2}| \ll t_k^{1/2} k^{-2}. \end{aligned} \tag{4.24}$$

We recall that $\{Y_k, k \geq 1\}$ is a sequence of independent multivariate normal random vectors with mean zero and covariance matrix Γ_d. Hence for any Brownian motion $\{X(t), t \geq 0\}$ in B with covariance structure given by T the sequences $\{(t_k - t_{k-1})^{-1/2} \Pi_{d_k}(X(t_k) - X(t_{k-1})), k \geq 1\}$ and $\{Y_k, k \geq 1\}$ have the same distribution. Hence by Kolmogorov's existence theorem, which remains valid in the Banach space setting, we can redefine the process $\{\sum_{\nu \leq t} x_\nu, t \geq 0\}$ and the sequence $\{Y_k, k \geq 1\}$ on a richer probability space without changing their joint distribution such that on this probability space there exists Brownian motion $\{X(t), t \geq 0\}$ with covariance structure given by T and satisfying

$$(t_k - t_{k-1})^{-1/2} \Pi_{d_k}(X(t_k) - X(t_{k-1})) = Y_k \qquad k \geq 1. \tag{4.25}$$

Indeed, if F denotes the joint distribution of finitely many X_k's and Y_k's and if G denotes the joint distribution of the same Y_k's and a finite number of properly normalized increments of the Brownian motion $\{X(t), t \geq 0\}$ then Lemma A1 of Berkes and Philipp (1979) shows that the consistency requirement in Kolmogorov's existence theorem is satisfied..

We now show that $\{X(t), t \geq 0\}$ has the desired properties. Let $\epsilon > 0$ and choose N so large that the conclusions of Lemmas 4.1 and 4.2 hold. Let $t > 0$ be given and define m by $t_m \leq t$

$\leq t_{m+1}$. Then

$$
\begin{aligned}
\| \textstyle\sum_{\nu\leq t} x_\nu - X(t) \| \leq & \| \textstyle\sum_{\nu\leq t} (x_\nu - \Pi_N x_\nu) \| \\
& + \| \Pi_N \| \max_{t_m < t \leq t_{m+1}} \| \textstyle\sum_{t_m<\nu\leq t} x_\nu \| \\
& + \| \Pi_N \| \textstyle\sum_{k\leq m} \| \textstyle\sum_{\nu\in I_k} x_\nu \| \\
& + \textstyle\sum_{k\leq m} \| \textstyle\sum_{\nu\in H_k} \Pi_N x_\nu - \Pi_N(X(t_k) - X(t_{k-1})) \| \\
& + \| \Pi_N \| \max_{t_m\leq t\leq t_{m+1}} \| X(t_m) - X(t) \| + \| \Pi_N X(t) - X(t) \| .
\end{aligned} \tag{4.26}
$$

By relation (2.3) of Kuelbs (1976a) we have $\| x \| < c \| x \|_{H_\mu}$ for $x \in H_\mu$ and by definition of $\| \cdot \|_{H_\mu}$ we have $\| y \|_{H_\mu} = | y |$ for $y \in \Pi_{d_k} B = \mathbb{R}^d$. Here $c \geq 1$ is a constant. Moreover, for $d_k \geq N$ we have $\Pi_N \circ \Pi_{d_k} = \Pi_N$ and $\| \Pi_N(\Pi_{d_k} z) \|_{H_\mu} \leq \| \Pi_{d_k} z \|_{H_\mu}$ for $z \in B$ since the Π_N's when restricted to H_μ are projections and hence have norm 1. Consequently, we have for $d_k \geq N$ by (4.4), (4.24), (4.25) and (4.14)

$$
\begin{aligned}
& \| \textstyle\sum_{\nu\in H_k} \Pi_N x_\nu - \Pi_N(X(t_k) - X(t_{k-1})) \| \\
& \quad = \| \Pi_N(\textstyle\sum_{\nu\in H_k} \Pi_{d_k} x_\nu - \Pi_{d_k}(X(t_k) - X(t_{k-1}))) \| \\
& \quad \leq c \| \Pi_N(\textstyle\sum_{\nu\in H_k} \Pi_{d_k} x_\nu - \Pi_{d_k}(X(t_k) - X(t_{k-1}))) \|_{H_\mu} \\
& \quad \leq c | \textstyle\sum_{\nu\in H_k} \xi_\nu - (t_k - t_{k-1})^{1/2} Y_k | \ll t_k^{1/2} k^{-2} \qquad \text{a.s.}
\end{aligned} \tag{4.27}
$$

After dividing (4.26) by $(t \log \log t)^{1/2}$ we obtain from Lemmas 4.1–4.3, 4.5, Proposition 4.1 and (4.27)

$$\limsup_{t\to\infty} (t \log \log t)^{-1/2} \| \textstyle\sum_{\nu\leq t} x_\nu - X(t) \| \leq 2\eta \qquad \text{a.s.}$$

since by the fact t_k increases we have

$$\textstyle\sum_{k\leq m} t_k^{1/2} k^{-2} \ll t_m^{1/2} \ll t^{1/2}.$$

4.4. *Proof of Theorem* 1. At first we prove a central limit theorem slightly stronger than Corollary 1.

PROPOSITION 4.2. *Let $\{x_j, j \geq 1\}$ be a weak sense stationary sequence of B-valued random variables with* $(2 + \delta)$th *moments with* $0 < \delta \leq 1$, *bounded by* 1 *and satisfying a strong mixing condition* (2.1) *with rate of decay* (2.2). *Moreover, suppose that for every* $\rho > 0$ *there exists a mapping* $\Lambda_\rho: B \to B$ *with finite dimensional range satisfying* (1.7)–(1.9). *Then the series defining* $T(f, g)$ *converges absolutely for all* $f, g \in B^*$. *Moreover,* $\{\mathscr{L}(n^{-1/2} S_n), n \geq 1\}$ *converges weakly to a mean zero Gaussian measure* μ *(say) whose covariance structure is given by* $T(f, g)$.

PROOF. Since the $(2 + \delta)$th moments are uniformly bounded by 1 we have for all $f \in B^*$

$$\sup_{\nu\geq 1} E | f(x_\nu) |^{2+\delta} \leq \| f \|^{2+\delta} < \infty. \tag{4.28}$$

Moreover, by Hölder's inequality and (1.9) we have for all $\rho > 0$

$$
\begin{aligned}
| E(f(x_\nu)) | & = | E(f(x_\nu)) - E\{f(\Lambda_\rho(x_\nu))\} | \\
& \leq \| f \| E \| x_\nu - \Lambda_\rho(x_\nu) \| \leq \| f \| \rho^{1/2}.
\end{aligned}
$$

Thus for all $f \in B^*$ and $\nu \geq 1$

$$E(f(x_\nu)) = 0. \tag{4.29}$$

Hence by (4.28), (2.2) and by (2.4) with $s = t = 2 + \delta$ and $r = 1 + 2/\delta$ we have for all $f, g \in B^*$ and $\nu \geq 1$

$$E\{f(x_\nu) g(x_1)\} \ll \nu^{-(1+\epsilon)}.$$

This proves the first claim. To prove the second claim we observe that by (1.9) and by

Theorem 4 the sequence $\{\mathscr{L}(n^{-1/2}S_n), n \geq 1\}$ of probability measures is relatively compact. By Proposition 2.1 the finite dimensional distributions of $\{n^{-1/2}S_n\}$ converge to mean zero Gaussian distributions determined by the covariance $T(f, g)$. Since a measure on B is determined by its finite dimensional distributions we conclude that the measures $\mathscr{L}(n^{1/2}S_n)$ converge weakly to a mean zero Gaussian measure on B with covariance $T(f, g)$. □

We apply Theorem 6 with the measure μ given by Proposition 4.2. From the way μ was obtained it is obvious that the hypothesis on the limit distribution of $\{\Pi_N x_{j+a}, j \geq 1\}$ is satisfied. In view of (1.9) Proposition 3.1 implies (4.2). Hence all that remains to be shown for the completion of the proof of Theorem 1 is the following lemma.

LEMMA 4.6. *With probability* 1

$$\lim_{n\to\infty} \| S_n/a_n - K \| = 0$$

where K is the unit ball of the Hilbert space determined by $T(f, g)$ (or μ) as in Lemma 2.1 *of Kuelbs* (1976*a*).

PROOF. We apply 3.1 of Kuelbs (1976a) to the sequence $\{S_n/a_n, n \geq 1\}$. *Let* $f \in B^*$. By (4.28) and (4.29) we see that Theorem 4 above holds for the sequence $\{f(x_j), j \geq 1\}$. But Theorem 4 implies the law of the iterated logarithm. Hence we have with probability 1

$$\limsup_{n\to\infty} f(S_n/a_n) = T^{1/2}(f, f).$$

Since $\{n^{-1/2}f(S_n), n \geq 1\}$ converges in distribution to μf^{-1} and since by Lemma 2.5 $\{n^{-1}(f(S_n))^2, n \geq 1\}$ is uniformly integrable we have that

$$\lim_{n\to\infty} n^{-1}E(f(S_n))^2 = \int_B x^2\, d\mu f^{-1}(x) = \int_B (f(x))^2\, d\mu(x).$$

But by Lemma 2.3

$$\lim_{n\to\infty} n^{-1}E(f(S_n))^2 = T(f, f).$$

Consequently we obtain from relation (2.5) of Kuelbs (1976a)

$$T^{1/2}(f, f) = \left(\int_B (f(x))^2\, d\mu(x)\right)^{1/2} = \sup_{x\in K} f(x).$$

Hence condition (3.1) of Theorem 3.1 of Kuelbs (1976a) is satisfied and the lemma will follow if we can show that the sequence $\{S_n/a_n, n \geq 1\}$ is with probability 1 relatively ($\equiv$ conditionally) compact. But this follows from a double barrelled application of Theorem 5 above. Given $\rho > 0$ we see that the sequence $\{\sum_{\nu\leq n} \Lambda_\rho(x_\nu)/a_n, n \geq 1\}$ is relatively compact with probability 1 by applying Theorem 5 to the finite dimensional random variables $\Lambda_\rho(x_\nu)$. But Theorem 5 also implies

$$\limsup_{n\to\infty} a_n^{-1} \| \textstyle\sum_{\nu\leq n} (x_\nu - \Lambda_\rho(x_\nu)) \| \leq 2000\rho^{1/2}/\delta \qquad \text{a.s.}$$

This together with the previous remarks implies the relative compactness of $\{S_n/a_n, n \geq 1\}$ and thus the lemma.

4.5. *Proof of Theorem* 2. Let

$$d_k = k^{\delta(1-\beta)/10}. \tag{4.30}$$

In view of (4.13) and because of our convention interpreting α as δ in case of ϕ-mixing random variables (4.30) is an admissible choice. All that is needed for the proof of Theorem 2 are the following two lemmas.

LEMMA 4.7. *As $k \to \infty$*

$$\| \textstyle\sum_{\nu\in H_k} (x_\nu - \Pi_{d_k} x_\nu) \| \ll t_k^{1/2} k^{-1} \qquad \text{a.s.}$$

PROOF. Since by (4.30) and (1.12)

$$E\|\Pi_{d_k}x_\nu - x_\nu\|^{2+\delta} \le 8(\|\Pi_{d_k}\|_1^{2+\delta} + 1)E\|x_\nu\|^{2+\delta} \ll e^{3d_k}$$

and since by (1.13) (recall that n_k is defined in the proof of Proposition 4.1)

$$b\sigma^{-2-\delta} \ll e^{3d_k} \cdot d_k^{36/\delta} \ll e^{4d_k} \ll 2^{n_k\delta/40}$$

we obtain from Proposition 3.1

$$P\{\|\sum_{\nu\in H_k}(x_\nu - \Pi_{d_k}x_\nu)\| \ge t_k^{1/2}k^{-1}\}$$

$$\ll P\{\|\sum_{\nu\in H_k}(x_\nu - \Pi_{d_k}x_\nu)\| \ge 10(10/\delta)^{(2+\delta)/\delta}\sigma_{d_k}(N_k \log\log N_k)^{1/2}\}$$

$$\ll N_k^{-2} + N_k^{-\delta/40} + n_k^{-11/\delta} \ll k^{-1.1}$$

since by (4.7), (4.9), (1.13), (4.30)

$$\sigma_{d_k}(N_k \log\log N_k)^{1/2} = o(t_k^{1/2}k^{-1}).$$

The lemma follows now from the Borel-Cantelli lemma.

LEMMA 4.8. *Let $\{X(t), t \ge 0\}$ be a Brownian motion with covariance structure T. Then as $k \to \infty$*

$$\|X(t_k) - X(t_{k-1}) - \Pi_{d_k}(X(t_k) - X(t_{k-1}))\| \ll t_k^{1/2}k^{-1} \qquad \text{a.s.}$$

For the proof of this lemma we shall apply Lemma 4.7 to the sequence $\{X(\nu + 1) - X(\nu), \nu \ge 1\}$. To see that Lemma 4.7 applies we need only verify that an analogue of (1.13) holds. That is, if $\mu = \mathscr{L}(X(1)) = \mathscr{L}(X(\nu + 1) - X(\nu))$ denotes the mean zero Gaussian limit of $\{n^{-1/2}S_n, n \ge 1\}$ then Lemma 4.8 is proved if we can show that there exists a constant C such that

$$\int_B \|x - \Pi_N x\|^2 \, d\mu(x) \le CN^{-12/\delta}. \tag{4.31}$$

For the proof of (4.31) we first show that $\{n^{-1}\|S_n\|^2, n \ge 1\}$ is uniformly integrable. Fix $\epsilon > 0$ and choose L so large that (1.13) implies

$$\sup_{n\ge1} n^{-1}E\|S_n - \Pi_L S_n\|^2 < \epsilon. \tag{4.32}$$

Since Π_L has finite-dimensional range we can apply Lemma 3.1 to each of the coordinates of $\Pi_L S_n$ and conclude that $\{n^{-1}\|\Pi_L S_n\|^2, n \ge 1\}$ is uniformly integrable. This together with (4.32) implies the uniform integrability of $\{n^{-1}\|S_n\|^2, n \ge 1\}$. Consequently since by Corollary 1 $\{n^{-1/2}S_n, n \ge 1\}$ converges in distribution to μ and since $I - \Pi_N$ is a linear operator on B we conclude that

$$\lim_{n\to\infty} n^{-1}E\|S_n - \Pi_N S_n\|^2 = \int_B \|x - \Pi_N x\|^2 \, d\mu(x).$$

(4.31) follows now from (1.13). □

We finally can finish the proof of Theorem 2. We have similar to (4.26) and (4.27)

$$\begin{aligned}\|\textstyle\sum_{\nu\le t}x_\nu - X(t)\| \le &\textstyle\sum_{k\le m}\|\sum_{\nu\in H_k}(x_\nu - \Pi_{d_k}x_\nu)\| \\ &+ \max_{t_m<t<t_{m+1}}\|\textstyle\sum_{t_m\le\nu\le t}x_\nu\| \\ &+ \max_{t_m<t<t_{m+1}}\|X(t) - X(t_m)\| \\ &+ \textstyle\sum_{k\le m}\|\sum_{\nu\in I_k}x_\nu\| \\ &+ \textstyle\sum_{k\le m}\|\sum_{\nu\in H_k}\Pi_{d_k}x_\nu - \Pi_{d_k}(X(t_k) - X(t_{k-1}))\|\end{aligned}$$

$$+ \sum_{k\le m} \| X(t_k) - X(t_{k-1}) - \Pi_{d_k}(X(t_k) - X(t_{k-1})) \|$$

$$\ll (t_m/\log t_m)^{1/2} \qquad \text{a.s.}$$

by Lemma 4.7, Proposition 4.1, Lemma 4.5, Lemma 4.3, (4.24), (4.25) and Lemma 4.8.

4.6. *Proof of Corollaries* 2, 3 *and* 4. We first observe that the standard lemmas on ϕ-mixing random variables as given in Billingsley (1968) pages 170–172 remain valid for ϕ-mixing random variables with values in a separable Hilbert space.

To prove Corollary 2 we apply Theorem 1 and Proposition 4.2. This can be done as soon as we prove the existence of the maps Λ_ρ satisfying (1.7)–(1.9). In view of Theorem 4 we can assume without loss of generality that H is infinite dimensional. We choose a complete orthonormal basis $\{e_i, i \ge 1\}$ for H and define the sequence of projections

$$L_k(x) = \sum_{i\le k}(x, e_i)e_i \qquad \text{for} \quad x \in H \qquad\qquad k \ge 1.$$

By (20.32) and Lemma 1, page 170 of Billingsley (1968)

$$E\| \textstyle\sum_{j\le n} (x_j - L_k(x_j)) \|^2 \le 4nE\| x_1 - L_k(x_1) \|^2 \textstyle\sum_{k\ge 1} \phi^{1/2}(k). \tag{4.33}$$

Since (1.6) holds $\sum_{k\ge 1} \phi^{1/2}(k) = A < \infty$. We now define for $\rho > 0$

$$\Lambda_\rho = L_{k(\rho)}$$

where

$$k(\rho) = \inf\{k : E\| x_1 - L_k(x_1) \|^2 \le \tfrac{1}{4}\rho/A\}.$$

By (4.33) Λ_ρ satisfies (1.7)–(1.9). Hence Corollary 2 is proved.

We now prove Corollary 3. In view of Theorem 4 we again can assume without loss of generality that H is infinite dimensional. We now make use of the remarks preceeding Corollary 3 relating the Gaussian measure μ and (1.15)–(1.17). Let Π_N be given as in (1.16). Of course, if M is finite-dimensional we have only finitely many Π_N's. We now define

$$x_j^* = (I - L)(x_j) = \sum_{i\ge 1} (x_j, e_i)e_i \qquad\qquad j \ge 1$$

where I is the identity map and where L is given by (1.17).

Recalling both expressions for Π_N in (1.16) we see that

$$\| \Pi_N \|_1^2 = \sup_{\|x\|\le 1} \| \Pi_N(x) \|_{H_\mu}^2$$

$$= \sup_{\|x\|\le 1} \sum_{i\le n}(x, \alpha_i)^2$$

$$= \sup_{\|x\|\le 1} \sum_{i\le N} \frac{(x, e_i)^2}{\lambda_i}$$

$$\le \sup_{1\le i\le N} \lambda_i^{-1}.$$

Since $\lambda_i = E(Z, e_i)^2$ and (1.18) holds we have (1.12). Thus to apply Theorem 2 to the sequence $\{x_j^*, j \ge 1\}$ we need only verify (1.13). However, arguing as in (4.33) we obtain

$$E\| \textstyle\sum_{j=a+1}^{a+n} (x_j^* - \Pi_N x_j^*) \|^2 \le 4nE\| x_1^* - \Pi_N(x_1^*) \|^2 \textstyle\sum_{k\ge 1} \phi^{1/2}(k)$$

$$\le 4n \textstyle\sum_{k\ge 1} \phi^{1/2}(k) \textstyle\sum_{i\ge N+1} E(x_1, e_1)^2$$

$$\le CnN^{-12/\delta}$$

for some constant C. Thus (1.14) holds for the sequence $\{x_j^*, j \ge 1\}$.

Hence to complete the proof of Corollary 3 it is enough to show that with probability 1

$$\| \textstyle\sum_{j\le t} (x_j - x_j^*)\| \ll t^{1/2}/\log t. \tag{4.34}$$

This will follow from the following proposition.

PROPOSITION 4.3. *Let $\{y_j, j \geq 1\}$ be a strict sense stationary sequence of ϕ-mixing random variables with values in a separable Hilbert space H. Assume that y_1 has mean zero and finite $(2 + \delta)$th moment with $0 < \delta \leq 1$. Suppose that (1.6) holds and that the covariance function $T(f, g)$ of the sequence $\{y_j, j \geq 1\}$ vanishes for all $f, g \in H^*$. Then for any $\gamma > 0$*

$$\|\sum_{j\leq t} y_j\| \ll t^\gamma \quad \text{a.s.}$$

Recall that the covariance function of a sequence of random variables is defined in (1.10).

For the proof of (4.34) we apply Proposition 4.3 to $y_j = x_j - x_j^* = L(x_j)$. We observe that the sequence $\{L(x_j), j \geq 1\}$ obeys the central limit theorem by Corollary 2 with limiting measure $\mu^L = \delta_0$, the unit mass at 0. Hence the covariance function T of the sequence $\{L(x_j), j \geq 1\}$ as defined in (1.10) vanishes identically.

We break up the proof of Proposition 4.3 into several steps which we formulate as lemmas, all valid under the hypothesis of the proposition.

LEMMA 4.9. $\sigma_y^2 \equiv \lim_{n\to\infty} n^{-1}E\|\sum_{j\leq n} y_j\|^2 = 0$.

PROOF. We fix $\epsilon > 0$ and choose a finite-dimensional projection R such that $E\|y_1 - R(y_1)\|^2 < \epsilon$. Let d denote the dimension of the range of R and $e_1, \cdots, e_d$ an orthonormal basis for $R(H)$. Then arguing as in (4.33)

$$n^{-1}E\|\sum_{j\leq n} (I - R)(y_j)\|^2 \leq C\epsilon \tag{4.35}$$

for some constant C. Furthermore.

$$n^{-1}E\|\sum_{j\leq n} R(y_j)\|^2 = n^{-1} \sum_{i,j\leq n} E\{(R(y_i), R(y_j))\}. \tag{4.36}$$

Now

$$\lim_{n\to\infty} n^{-1} \sum_{i,j\leq n} E\{(R(y_i), e_r)(R(y_j), e_r)\} = T(R^*(e_r), R^*(e_r)) = 0$$

for $r = 1, \cdots, d$. Here R^* is the adjoint of R. Hence by (4.36)

$$n^{-1}E\|\sum_{j\leq n} R(y_j)\|^2 < \epsilon \tag{4.37}$$

for sufficiently large n. The result follows now from (4.35), (4.37) and Minkowski's inequality.

LEMMA 4.10. $E\|\sum_{j\leq n} y_j\|^2 \ll 1$.

PROOF. This follows, for instance, from Billingsley (1968), page 172 as $\phi(n) \ll n^{-6}$ by (1.6).

We now can finish the proof of Proposition 4.3 by applying the Gaal-Koksma strong law of large numbers which remains valid for random variables with values in a linear metric space. (See Theorem A1 in Philipp and Stout (1975), page 134.)

Hence Corollary 3 is proved.

Finally Corollary 4 follows easily by combining Corollary 3 and Theorem 2.4 of Kuelbs (1975b). That is, if $\psi(t)$ is as in Corollary 4, then standard methods imply that it suffices to show the equivalence of (1.22) and (1.23) only for those $\psi(t)$ which also satisfy

$$(\log \log t)^{1/2} \leq \psi(t) \leq 2(\log \log t)^{1/2}$$

for large t. For this class of functions ψ Corollary 3 and Theorem 2.4 of Kuelbs (1975b) easily yield Corollary 4.

5. Proof of Theorem 3. Suppose that x_1 is pre-Gaussian and let μ denote the mean zero Gaussian measure determined by the covariance structure

$$T(f, g) = E\{f(x_1)g(x_1)\} \qquad f, g \in B^*. \tag{5.1}$$

Further, assume that any of the conditions in (b) hold. Then by Theorem 4.1 of Kuelbs (1977) we have that all the conditions in (b) hold. In particular, condition (b-iii) implies that with probability 1

$$\lim_{n\to\infty} \| S_n/a_n - K \| = 0 \tag{5.2}$$

where K is the unit ball of the Hilbert space determined by μ. This shows that condition (4.3) in Theorem 6 is satisfied.

We put $d = \min(N, \dim H_\mu)$. Then as in Section 4.4 $\Pi_N B = \Pi_N H_\mu$ can be viewed as a Euclidean space $\mathbb{R}^d$ with metric induced by the norm of H_μ. Again we denote this metric by $|\cdot|$.

To verify the hypothesis on the limit distribution of $\{\Pi_N x_{j+a}, j \ge 1\}$ we observe that it is a sequence of independent identically distributed $\mathbb{R}^d$-valued random variables, centered at expectations, with finite $(2+\delta)$th moments and with Γ_d, the $d \times d$ identity matrix, as their covariance matrix. Hence the sequence satisfies the central limit theorem. The limiting Gaussian distribution λ_N also has Γ_d as its covariance matrix. Since the Gaussian measure μ has covariance structure given by (5.1) we conclude that μ^{Π_N} is also a Gaussian measure with covariance matrix Γ_d when $\Pi_N B$ is interpreted as above. Since all measures have mean zero we therefore have $\lambda_N = \mu^{\Pi_N}$.

It remains to verify (4.2). We prove a slightly stronger result.

PROPOSITION 5.1. *Let $\{x_j, j \ge 1\}$ be a sequence of independent identically distributed B-valued random variables centered at expectations and $(2+\delta)$th moments bounded by* 1. *Suppose that condition* (b)(ii) *in Theorem* 3 *holds. Then there exists a constant C such that for all R, n* ≥ 1

$$P\{\| \textstyle\sum_{j\le n} x_j \| \ge 9R(n \log\log n)^{1/2}\} \le C(R^{-2}n^{-\delta/4} + \exp(-1.1R^{\delta/(2+\delta)}\log\log n)).$$

The proof consists of a simplification of the proof of Proposition 3.1. We put

$$\begin{aligned} w_j &= x_j \qquad \text{if} \quad \| x_j \| \le R^{2/(2+\delta)}(n/\log\log n)^{1/2} \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

Then we obtain from Markov's inequality

$$P\{w_j \ne x_j\} \ll R^{-2}n^{-1-\delta/4}. \tag{5.3}$$

LEMMA 5.1. $E\| \sum_{j\le n} w_j \| = o((n \log\log n)^{1/2})$.

PROOF. We introduce the event

$$C = \cup_{1\le \nu \le n} \cup_{1\le j_1 < \cdots < j_\nu \le n} C(j_1, \cdots, j_\nu)$$

where

$$C(j_1, \cdots, j_\nu) = \{w_{j_1} \ne x_{j_1}, \cdots, w_{j_\nu} \ne x_{j_\nu}, w_j = x_j \quad j \ne j_i \quad (1 \le i \le \nu)\}.$$

Then by (5.3)

$$P\{C(j_1, \cdots, j_\nu)\} \ll n^{-(1+\delta/4)\nu}$$

and

$$\begin{aligned} &\int_{C(j_1,\cdots,j_\nu)} \| x_{j_i} \| \\ &\le P\{w_{j_l} \ne Y_{j_l}, 1 \le l \le \nu, l \ne i; w_j = Y_j \quad \text{for all} \quad j \ne j_l, 1 \le l \le \nu\} \int_{\{w_{j_i} \ne x_{j_i}\}} \| x_{j_i} \| \\ &\ll n^{-(1+\delta/4)(\nu-1)} n^{-1/2-\delta/8}. \end{aligned}$$

On $C(j_1, \cdots, j_\nu)$ we have

$$\|\textstyle\sum_{j\leq n} w_j\| \leq \|\sum_{j\leq n} x_j\| + \sum_{i\leq \nu} \|x_{j_i}\|.$$

Hence

$$\int_C \|\textstyle\sum_{j\leq n} w_j\| \leq \int_C \|\sum_{j\leq n} x_j\| + \sum_{\nu\leq n} n^\nu \nu n^{-(1+\delta/4)(\nu-1/2)}$$
$$= \int_C \|\textstyle\sum_{j\leq n} x_j\| + O(n^{1/2}).$$

On C^c the two sums are identical. Hence we obtain the lemma from (b)(ii). □

We now can finish the proof of Proposition 5.1. By Lemma 3.4 with $k = n$, $b_k = n^{1/2}$, $c = R^{2/(2+\delta)}(\log\log n)^{-1/2}$ and $\epsilon = \frac{9}{2}R(\log\log n)^{1/2}$ we have

$$P\{\|\textstyle\sum_{j\leq n} w_j\| \geq 9R(n\log\log n)^{1/2}\} \leq \exp(-\frac{9}{2}R^{\delta/(2+\delta)}\log\log n(1 - \frac{3}{4}) + o(\log\log n))$$
$$\ll \exp(-1.1R^{\delta/(2+\delta)}\log\log n).$$

Hence by (5.3)

$$P\{\|\textstyle\sum_{j\leq n} x_j\| \geq 9R(n\log\log n)^{1/2}\}$$
$$\leq \textstyle\sum_{j\leq n} P\{w_j \neq x_j\} + P\{\|\sum_{j\leq n} w_j\| \geq 9R(n\log\log n)^{1/2}\}$$
$$\ll R^{-2}n^{-\delta/4} + \exp(-1.1R^{\delta/(2+\delta)}\log\log n).$$

This concludes the proof of Proposition 5.1 and of half of Theorem 3.

To prove the second half we assume that (1.11) holds. Then S_n/a_n converges to zero in probability since $X(n)/a_n$ does. Hence in view of Theorem 4.1 of Kuelbs (1977) we conclude that conditions (i), (ii) and (iii) hold. Thus it suffices to show that x_1 is pre-Gaussian.

Let $f \in B^*$. Then, by the classical law of the iterated logarithm we have that with probability 1

$$\limsup_{n\to\infty} f(S_n)/a_n = E^{1/2}\{f^2(x_1)\} = T^{1/2}(f,f)$$

and

$$\limsup_{n\to\infty} f(X(n))/a_n = E^{1/2}\{f^2(X(1))\}.$$

Hence by (1.11)

$$T(f,f) = E\{f^2(X(1))\}. \tag{5.4}$$

Since real symmetric bilinear forms are determined by their diagonals (5.4) implies

$$E\{f(x_1)g(x_1)\} = T(f,g) = E\{f(X(1))g(X(1))\} \qquad f,g\in B^*.$$

Hence x_1 is pre-Gaussian by definition. □

We shall discuss now the example due to Kuelbs (1976b) and mentioned at the end of Section 1. Let $x_1, x_2, \cdots$ be the independent identically distributed c_0-valued random variables defined by

$$x_\nu = \textstyle\sum_{j=1}^\infty \epsilon_j^{(\nu)} e_j(2\log j)^{-1/2} \qquad \nu\geq 1$$

where $\{e_j, j\geq 1\}$ is the canonical basis of c_0 and $\{\epsilon_j^{(\nu)}, j\geq 1\}$ are independent sequences of independent identically distributed random variables with $P\{\epsilon_j^{(\nu)} = \pm 1\} = \frac{1}{2}$. Then as is shown in Kuelbs (1976b) the sequence $\{x_\nu, \nu\geq 1\}$ satisfies condition (iii) of Theorem 3. Hence by Theorem 4.1 of Kuelbs (1977) conditions (i) and (ii) of Theorem 3 are also satisfied. But it is impossible to have a Brownian motion $\{X(t), t\geq 0\}$ satisfying (1.11) since x_1 is not pre-Gaussian.

On the other hand, for each $\rho > 0$ we can redefine the sequence $\{x_\nu, \nu\geq 1\}$ on a new probability space together with a Brownian motion $\{X_\rho(t), t\geq 0\}$ such that (4.1) holds. To prove this claim we use the fact that $\{x_\nu,\ \nu \geq 1\}$ satisfies a compact law of the iterated logarithm (see Kuelbs (1976b) for details) with limit set

$$K = \{\{\alpha_k,\ k \geq 1\} \in c_0\colon \textstyle\sum_{k\geq 1} \alpha_k^2 \log k \leq ½\}. \tag{5.5}$$

Define now for $N \geq 1$

$$\psi_N x_\nu = \textstyle\sum_{j\leq N} \epsilon_j^{(\nu)} e_j (2 \log j)^{-1/2}.$$

Since the map $I - \psi_N$ is continuous (5.5) implies that for any $\rho > 0$ there is an N such that with probability 1 (see the proof of Lemma 4.1)

$$\limsup_{n\to\infty} a_n^{-1} \| \textstyle\sum_{\nu\leq n} x_\nu - \psi_N x_\nu \| \leq \rho. \tag{5.6}$$

Using Theorem 4 we can construct an N-dimensional Brownian motion $\{X_\rho(t),\ t \geq 0\}$ such that with probability 1

$$\| \textstyle\sum_{\nu\leq t} \psi_N x_\nu - X_\rho(t) \| \ll t^{1/2-\lambda} \tag{5.7}$$

for some $\lambda > 0$. Combining (5.6) and (5.7) we have (4.1) as claimed.

This example shows that approximating partial sum processes of B-valued random variables by the corresponding B-valued Brownian motion in the form of (4.1) is intrinsically somewhat less precise than one might hope. More specifically, the above sequence $\{x_\nu,\ \nu \geq 1\}$ satisfies the compact law of the iterated logarithm (and hence also the functional law of the iterated logarithm by Kuelbs (1975a)) yet it cannot be approximated by a Brownian motion so that (1.11) holds. Of course, the problem in this example is that there is no natural limiting Brownian motion to approximate with.

6. An application to the uniform law of the iterated logarithm for classes of functions. Kaufman and Philipp (1978) proved, among a number of things, the following uniform law of the iterated logarithm. Let S_α be the class of real-valued functions on [0, 1] with $f(0) = f(1)$, ;$\int_0^1 f(x)\,dx = 0$ and satisfying the Lipschitz condition

$$|f(x) - f(y)| \leq |x - y|^\alpha \qquad 0 \leq x, y \leq 1. \tag{6.1}$$

If $\alpha > ½$ and $\{\xi_j, j \geq 1\}$ is either a strictly stationary sequence of random variables, uniformly distributed over [0, 1] and satisfying a strong mixing condition or a sequence of lacunary random variables, then with probability 1

$$\limsup_{n\to\infty} a_n^{-1} \sup_{f\in S_\alpha} | \textstyle\sum_{\nu\leq n} f(\xi_\nu) | \leq C \tag{6.2}$$

where C is some finite constant. Moreover, they proved that if $\alpha < ½$ than (6.2) is false even if $\{\xi_j, j \geq 1\}$ is a sequence of independent random variables uniformly distributed over [0, 1].

In this section we prove an almost sure invariance principle for $\alpha > ½$ and stationary, ϕ-mixing sequences of random variables $\{\xi_j,\ j \geq 1\}$. In view of the remarks in Section 1, this result will, of course, imply (6.2).

As a matter of fact we shall prove a theorem for a class slightly bigger than S_α. Let $\lambda = \{\lambda_n,\ -\infty < n < \infty\}$ be any fixed sequence of positive numbers such that

$$C_\lambda =_{\text{def}} \textstyle\sum_{|n|\geq 0} \lambda_n^{-2} < \infty. \tag{6.3}$$

Let S be the class of all real-valued continuous functions f on [0, 1] with $\int_0^1 f(x)\,dx = 0$ and whose Fourier series

$$f(x) = \textstyle\sum_{|n|\geq 1} c_n e^{2\pi i n x} \tag{6.4}$$

is such that the Fourier coefficients $\{c_n\}$ satisfy the condition

$$\textstyle\sum_{|n|\geq 0} |c_n|^2 \lambda_n^2 \leq 1. \tag{6.5}$$

Then $S \supseteq \cup_{\alpha>1/2} S_\alpha$ by taking $\lambda_n = |n|^{1/2} (\log |n|)^2$ (see Zygmund (1935), page 136, (3)). Hence Theorem 7 below will imply, as a simple corollary, (6.2) with S_α replaced by S.

The conditions (6.3) and (6.5) easily imply (6.4) converges absolutely and uniformly in $f \in S$. (See the proof of Lemma 6.1). Since $S \subset C[0, 1]$ we have a natural distance given by the uniform norm

$$\|f - g\|_\infty = \sup_{0 \le x \le 1} |f(x) - g(x)|.$$

Our theorem can now be stated as follows.

THEOREM 7. *Let $C(S)$ denote the class of real-valued continuous functions on S and let $\{\xi_j, j \ge 1\}$ be a ϕ-mixing sequence of random variables, uniformly distributed over* $[0, 1]$. *Assume that* (1.6) *holds. Define the random variables with values in $C(S)$*

$$x_j(f, \omega) = f(\xi_j(\omega)) \qquad f \in S, j \ge 1. \tag{6.6}$$

Suppose that $\{x_j, j \ge 1\}$ is weak sense stationary. Then $\{n^{-1/2} S_n, n \ge 1\}$ converges weakly to a mean zero Gaussian measure on $C(S)$. Moreover, we can redefine $\{\xi_j, j \ge 1\}$ on a new probability space on which there exists Brownian motion with covariance given by (1.10) such that

$$\| \textstyle\sum_{j \le t} x_j - X(t) \| = o((t \log\log t)^{1/2}) \quad \text{a.s.} \tag{6.7}$$

REMARKS. 1. The norm on $C(S)$ is, of course, that given by the supremum norm

$$\|y\| = \sup_{f \in S} |y(f)| \qquad y \in C(S). \tag{6.8}$$

Since, as will be shown in Lemma 6.1 below, S is a compact metric space it will follow that $C(S)$ is a separable Banach space.

2. If $\{\xi_j, j \ge 1\}$ is strictly stationary, then the $C(S)$-valued random variables are also strictly stationary.

6.1. *The compactness of S.*

LEMMA 6.1. *S is a compact metric space in the distance given by the uniform norm*

$$\|f - g\|_\infty = \sup_{0 \le x \le 1} |f(x) - g(x)|.$$

PROOF. We will first show that S is a uniformly bounded, equicontinuous subset of $C[0, 1]$. Then the Arzela-Ascoli theorem will imply that S is relatively compact.

To see that S is uniformly bounded we note that the Fourier coefficients of $f \in S$ satisfy $c_0 = \int_0^1 f(x)\, dx = 0$ and

$$\textstyle\sum_{|n| \ge 1} |c_n| \le \{\sum_{|n| \ge 1} |c_n|^2 \lambda_n^2\}^{1/2} \{\sum_{|n| \ge 1} \lambda_n^{-2}\}^{1/2} \le C_\lambda^{1/2} < \infty \tag{6.9}$$

by (6.3) and (6.5). Hence the Fourier series for each $f \in S$ converges absolutely and uniformly to f. Hence by (6.4) and (6.9)

$$\sup_{f \in S} \|f\|_\infty \le C_\lambda^{1/2} < \infty. \tag{6.10}$$

This proves that S is uniformly bounded.

We now show that S is equicontinuous. For each $\epsilon > 0$ there is N such that

$$\textstyle\sum_{|n| > N} \lambda_n^{-2} < \epsilon^2/8. \tag{6.11}$$

Let $\{c_n(f)\}$ denote the Fourier coefficients of $f \in S$. Then by (6.11), (6.3) and (6.5)

$$\begin{aligned} \sup_{f \in S} |f(x) - f(y)|^2 &= \sup_{f \in S} \left| \textstyle\sum_{|n| \ge 1} \lambda_n c_n(f) \cdot \dfrac{e^{2\pi i n x} - e^{2\pi i n y}}{\lambda_n} \right|^2 \\ &\le \sup_{f \in S} \textstyle\sum_{1 \le |n| \le N} \lambda_n^2 |c_n(f)|^2 \cdot \sum_{1 \le |n| \le N} \lambda_n^{-2} |e^{2\pi i n x} - e^{2\pi i n y}|^2 \\ &\quad + \sup_{f \in S} \textstyle\sum_{|n| > N} \lambda_n^2 |c_n(f)|^2 \cdot \sum_{|n| > N} 4\lambda_n^{-2} \\ &\le 4\pi^2 C_\lambda \cdot N^2 |x - y|^2 + \tfrac{1}{2}\epsilon^2 < \epsilon^2 \end{aligned}$$

if $|x-y| \le \frac{1}{10} C_\lambda^{-1/2} N^{-1} \epsilon = \delta$ (say).

Hence S is relatively compact. The compactness of S will follow if we show that S is closed. Let $\{f_k, k \ge 1\}$ be a sequence of elements of S converging uniformly to f. Then f is continuous, $\int_0^1 f(x)\,dx = 0$, and $\{f_k, k \ge 1\}$ converges to f in $L^2[0, 1]$. Hence the Fourier coefficients of f_k, call them $\{c_n^{(k)}, n \in \mathbb{Z}\}$ converge to the Fourier coefficients $\{c_n, n \in \mathbb{Z}\}$ of f. Since for all $k \ge 1$

$$\sum_{|n|\ge 1} |c_n^{(k)}|^2 \lambda_n^2 \le 1$$

we have for each $N \ge 1$

$$\sum_{1\le|n|\le N} |c_n|^2 \lambda_n^2 = \lim_{k\to\infty} \sum_{1\le|n|\le N} |c_n^{(k)}|^2 \lambda_n^2 \le 1.$$

Letting $N \to \infty$ we conclude that $f \in S$ and that thus S is closed.

6.2. *The maps* Λ_ρ. First we observe that the x_j are $C(S)$ valued random variables which, by (6.10), are uniformly bounded. Further, since $\int_0^1 f(x)\,dx = 0$, we have the Bochner integral of each $x_j (j \ge 1)$ equal to zero in $C(S)$. The x_j are ϕ-mixing with ϕ as in (1.6) since the ξ_j have this property. Hence to show that Theorem 1 is applicable we must obtain the maps $\Lambda_\rho : C(S) \to C(S)$ satisfying (1.7)–(1.9).

For this purpose we define

$$M = \cup_{0\le t\le 1} \{ y \in C(S); y(f) = f(t) \quad \text{for all} \quad f \in S\}. \tag{6.12}$$

Note that the mapping τ from M into $[0, 1]$, defined by $t = \tau(y)$ is continuous and that M is a closed subset of $C(S)$. Now define the mappings Ψ_N from $C(S)$ into $C(S)$ by

$$\begin{aligned} \Psi_N(y)(f) &= \mathrm{Re}\left(\sum_{|k|<N} c_k(f) \exp\{2\pi i k[N^2 \tau(y)]/N^2\}\right) && y \in M \\ &= 0 && \text{otherwise.} \end{aligned} \tag{6.13}$$

Here $[s]$ is the greatest integer in s. Also define the mappings Φ_N by

$$\begin{aligned} \Phi_N(y)(f) &= \mathrm{Re}\left(\sum_{|k|<N} c_k(f) \exp\{2\pi i k \tau(y)\}\right) && y \in M \\ &= 0 && \text{otherwise.} \end{aligned} \tag{6.14}$$

We observe that Ψ_N maps $C(S)$ onto the finite dimensional subspace generated by

$$\{ y_{kl}(\cdot): y_{kl}(f) = \mathrm{Re}(c_k(f))\cos(2\pi kl/N^2) \text{ or } \mathrm{Im}(c_k(f))\sin(2\pi kl/N^2), |k| < N, 0 \le l \le N^2\}.$$

Further it is easy to see that $\Psi_N(x_j)$ and $\Phi_N(x_j)$ are measurable for $j \ge 1$ and $N \ge 1$.

Condition (1.9) will follow from the following lemma.

LEMMA 6.2. *For each* $\rho > 0$ *there exists an* $N(\rho)$ *such that*

$$E\left\| \sum_{j=m+1}^{m+n} x_j - \Psi_N(x_j) \right\|^2 \le n\rho$$

for all $m \ge 0$, $n \ge 1$ *and* $N \ge N(\rho)$.

PROOF. We prove the lemma for $m = 0$. The general case follows applying this special case to the sequence $\{\xi_{j+m}, j \ge 1\}$. We first note that suppressing the index N in Φ and Ψ

$$E\| S_n - \sum_{j\le n} \Psi(x_j)\|^2 \le 2\{E\| S_n - \sum_{j\le n} \Phi(x_j)\|^2 + E\|\sum_{j\le n} \Phi(x_j) - \Psi(x_j)\|^2\}. \tag{6.15}$$

Now by (6.5), (6.6) and (6.14).

$$\begin{aligned} E\| S_n - \sum_{j\le n} \Phi(x_j)\|^2 & \\ &= E\{\sup_{f\in S} |\sum_{j\le n} \mathrm{Re}(\sum_{|k|\ge N} c_k(f)\exp(2\pi i k\xi_j))|^2\} \\ &\le E[\sup_{f\in S} (\sum_{|k|\ge N} |c_k(f)|^2 \lambda_k^2 \cdot \sum_{|k|\ge N} \lambda_k^{-2} |\sum_{j\le n} \exp(2\pi i k \xi_j)|^2)\} \\ &\le \sum_{|k|\ge N} \lambda_k^{-2} E|\sum_{j\le n} \exp(2\pi i k\xi_j)|^2 \\ &\le C_1 n \sum_{|k|\ge N} \lambda_k^{-2}. \end{aligned} \tag{6.16}$$

The last step follows from the proof of Lemma 2.3. Indeed, it is easy to show by the same standard arguments used to prove Lemma 2.3 that

$$E\{\textstyle\sum_{j\leq n}\cos(2\pi k\xi_j)\}^2 \leq \tfrac{1}{4}C_1 n \tag{6.17}$$

for some constant C_1 since ξ_j has uniform distribution over [0, 1]. A similar estimate holds if in (6.17) cosine is replaced by sine.

By (6.13)

$$\Psi(x_j)(f) = \mathrm{Re}(\textstyle\sum_{|k|<N} c_k(f)\exp(2\pi ik[N^2\xi_j]/N^2)). \tag{6.18}$$

Put

$$\eta_j = \eta_j(k) = \exp(2\pi ik\xi_j) - \exp(2\pi ik[N^2\xi_j]/N^2).$$

Then for $N \geq 1$

$$E\eta_j = 0 \qquad j \geq 1$$

and

$$|\eta_j| \leq 2\pi|k|N^{-2} \qquad j \geq 1.$$

Thus by the argument leading to (6.17)

$$E|\textstyle\sum_{j\leq n}\eta_j|^2 \leq C_1 n|k|^2 N^{-4}.$$

Consequently, by (6.18)

$$\begin{aligned} E\|\textstyle\sum_{j\leq n}\Phi(x_j) - \Psi(x_j)\|^2 &= E\{\sup_{f\in S}|\textstyle\sum_{j\leq n}\mathrm{Re}\sum_{|k|<N} c_k(f)\eta_j|^2\} \\ &\leq E\{\sup_{f\in S}\textstyle\sum_{|k|<N}|c_k(f)|^2\lambda_k^2\cdot\sum_{|k|<N}\lambda_k^{-2}|\sum_{j\leq n}\eta_j|^2\} \\ &\leq \textstyle\sum_{|k|<N}\lambda_k^{-2}C_1 n|k|^2N^{-4} \leq C_1C_\lambda nN^{-2}. \end{aligned} \tag{6.19}$$

The lemma follows now from (6.15), (6.16) and (6.19).

6.3. *Proof of Theorem* 7. As observed after the statement of Lemma 6.2 the mapping $\Lambda\rho = \Psi_{N(\rho)}$ will satisfy (1.9). Furthermore, we conclude from (6.9) and (6.18) that $\|\Lambda_\rho(x_j)\|$ is uniformly bounded. Finally for every $f \in S$ we have $E\Lambda_\rho(x_j)(f) = 0$. Hence Theorem 7 follows from Theorem 1.

Acknowledgement. We thank Professor William Pruitt for his careful reading of the manuscript and for his great many comments which helped to improve the presentation of the material immensely. We also thank Herold Dehling for his comments on the manuscript, in particular for an observation that gave Lemma 3.1 its present form.

REFERENCES

DE ACOSTA, ALEJANDRO and GINÉ, EVARIST (1979). Convergence of moments and related functionals in the general central limit problem. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **48** 213–231.

BERKES, ISTVÁN and PHILIPP, WALTER (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probability* **7** 29–54.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

DAVYDOV, YU. A. (1970). The invariance principle for stationary processes. *Theo. Probability Appl.* **15** 487–498.

DEHLING, HEROLD (1980). Limit theorems for sums of weakly dependent Banach space valued random variables. Unpublished manuscript.

DEO, CHANDRAKANT M. (1973). A note on empirical processes of strong mixing sequences. *Ann. Probability* **1** 870–875.

DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

DVORETZKY, ARYEH (1970). Asymptotic normality for sums of dependent random variables. *Proc. Sixth Berkeley Symposium* **II** 513–535.

FERNIQUE, X. (1970). Integrabilité des vecteurs Gaussiens. *C. R. Acad. Science Paris Ser. A.* **270** 1698–1699.

GAAL, S. and GAAL, L. (1964). The discrepancy of the sequence $\{(2^n x)\}$. *Proc. Koninkl. Nederl. Akad. Wetensch. Ser. A* **67** 129–143.

IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theo. Probability Appl.* **7** 349–382.

IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables.* Noordhoff, Groningen.

JAIN, N. and KALLIANPUR, G. (1970). Norm convergent expansions for Gaussian processes in Banach spaces. *Proc. Amer. Math. Soc.* **25** 890–895.

KAUFMAN, R. and PHILIPP, WALTER (1978). A uniform law of the iterated logarithm for classes of functions. *Ann. Probability* **6** 930–952.

KUELBS, J. (1975a). The law of the iterated logarithm and related strong convergence theorems for Banach space valued random variables. *Lecture Notes in Math.* Vol. 539 Springer, Berlin.

KUELBS, J. (1975b). Sample path behavior for Brownian motion in Banach spaces. *Ann. Probability* **3** 247–261.

KUELBS, J. (1976a). A strong convergence theorem for Banach space valued random variables. *Ann. Probability* **4** 744–771.

KUELBS, J. (1976b). A counterexample for Banach space valued random variables. *Ann. Probability* **4** 684–689.

KUELBS, J. (1977). Kolmogorov's law of the iterated logarithm for Banach space valued random variables. *Illinois J. Math.* **21** 784–800.

KUELBS, J. and LEPAGE, R. (1973). The law of the iterated logarithm for Brownian motion in a Banach space. *Trans. Amer. Math. Soc.* **185** 253–264.

PHILIPP, WALTER (1979). Almost sure invariance principles for sums of B-valued random variables. In *Probability in Banach Spaces II, Lecture Notes in Math.* **709** 171–193, Springer, New York.

PHILIPP, WALTER and STOUT, WILLIAM (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **161** 1–140.

PISIER, G. (1975). Le théorème de la limite centrale et la loi du logarithme iteré dans les espaces de Banach. In *Seminaire Maurey-Schwartz* 1975–76, *Expose* IV. L'Ecole Polytechnique, Paris.

ROZANOV, YU. A. and VOLKONSKII, V. A. (1959). Some limit theorems for random functions. I. *Theo. Probability Appl.* **4** 178–197.

SERFLING, R. J. (1968). Contributions to central limit theory for dependent variables. *Ann. Math. Statist.* **39** 1158–1175.

SOTRES, DAVID A. and MALAY GHOSH (1977). Strong convergence of linear rank statistics for mixing processes. *Sankhyā Ser B* **39** 1–11.

STRASSEN, V. (1964). An almost sure invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebeite* **3** 211–226.

ZYGMUND, A. (1935). *Trigonometrical Series.* Lwow, Warszawa.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WI 53706

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, IL 61801