Z. Wahrscheinlichkeitstheorie verw. Gebiete
63, 393–432 (1983)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Limit Theorems for Sums of Weakly Dependent Banach Space Valued Random Variables

Herold Dehling*

Institut für Mathematische Stochastik der Universität, Lotzestr. 13, 3400 Göttingen, Federal Republic of Germany

Summary. We prove an estimate for the Prohorov-distance in the central limit theorem for strong mixing Banach space valued random variables. Using a recent variant of an approximation theorem of Berkes and Philipp (1979) we obtain as a corollary a strong invariance principle for absolutely regular sequences with error term $t^{\frac{1}{2}-\gamma}$. For strong mixing sequences we prove a strong invariance principle with error term $o((t\log\log t)^{\frac{1}{2}})$.

1. Introduction

In a recent paper Kuelbs and Philipp (1980) proved several almost sure invariance principles for sums of ϕ-mixing random variables with values in a separable Banach space. As remarked in a paper of Dehling and Philipp (1982) [6] their theorems hold even for partial sums of absolutely regular sequences. One of the purposes of this paper is to improve some of the results of Kuelbs and Philipp (1980) even further by using a different approach. We establish first estimates for the Prohorov distance in the central limit theorem for strong mixing B-valued random variables. These results, which may be of independent interest, together with a recent variant (see [6]) of an approximation theorem of Berkes and Philipp (1979), yield then almost sure invariance principles.

First we introduce some notation. Throughout the paper B denotes a real separable Banach space. Let $\{X_\nu,\ \nu\geqq 1\}$ be a sequence of B-valued random variables and denote by $\mathfrak{M}_a^b$ the σ-field generated by the random variables $X_a, X_{a+1}, \dots, X_b$.

i) $\{X_\nu,\ \nu\geqq 1\}$ is said to satisfy a strong mixing condition if there exists a sequence of real numbers $\alpha(n)\downarrow 0$ such that

$$|P(A\cap B)-P(A)\,P(B)|\leqq\alpha(n) \tag{1.1}$$

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

* Supported by the Studienstiftung des deutschen Volkes

ii) $\{X_\nu, \nu \geqq 1\}$ is called absolutely regular if for some $\beta(n) \downarrow 0$

$$E \sup_{A \in \mathfrak{M}_{k+n}^{\infty}} |P(A|\mathfrak{M}_1^k) - P(A)| \leqq \beta(n) \tag{1.2}$$

for all $k, n \geqq 1$.

iii) $\{X_\nu, \nu \geqq 1\}$ is called ϕ-mixing if for some $\phi(n) \downarrow 0$

$$|P(A \cap B) - P(A)\,P(B)| \leqq \phi(n)\,P(A) \tag{1.3}$$

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

It is well known that ϕ-mixing implies absolutely regular which in turn implies strong mixing. Moreover α, β and ϕ can be chosen in such a way that

$$\alpha(n) \leqq \beta(n) \leqq \phi(n). \tag{1.4}$$

On the other hand there exist examples of stationary sequences that are strong mixing but not absolutely regular and there exist examples that are absolutely regular but not ϕ-mixing. For a detailed investigation of these properties for stationary Gaussian processes see Ibragimov and Rozanov (1977). From their results it is also easy to derive the examples mentioned above. We denote the space of all probability measures on B by $\mathcal{M}(B)$. If $\mu, \nu \in \mathcal{M}(B)$, their Prohorov-distance is defined by

$$\pi(\mu, \nu) = \inf\{\varepsilon > 0 \,|\, \mu(A) \leqq \nu(A^\varepsilon) + \varepsilon \text{ for all closed sets } A \subset B\}$$

where $A^\varepsilon = \{x \in B \,|\, \|x - A\| < \varepsilon\}$.

A measure μ on B is called Gaussian, if $\mu \circ f^{-1}$ is a normal distribution on the real line for all $f \in B^*$, the topological dual space of B. It is well-known that each Gaussian measure is completely determined by its mean and its covariance function T which is defined by

$$T(f, g) = \int f(x)\, g(x)\, d\mu(x), \qquad f, g \in B^*.$$

We denote therefore the Gaussian measure with mean a and covariance operator T by $N(a, T)$.

B-valued Brownian motion is defined in exactly the same way as Brownian motion on the real line. It is well-known that for each mean zero Gaussian measure μ on B there exists B-valued Brownian motion such that $X(1)$ has distribution μ.

Theorem A. *Let $\{B_k, m_k, k \geqq 1\}$ be a sequence of Polish spaces. Let $\mathfrak{B}_k$ denote the Borel field of B_k, let $\{X_k, k \geqq 1\}$ be a sequence of random variables with values in B_k and let $\mathfrak{F}_k$ be a non-decreasing sequence of σ-fields such that X_k is $\mathfrak{F}_k$-measurable. Suppose that for some sequence $\{\beta_k, k \geqq 1\}$ of non-negative numbers*

$$E \sup_{A \in \mathfrak{B}_k} |P(X_k \in A | \mathfrak{F}_{k-1}) - P(X_k \in A)| \leqq \beta_k \tag{1.5}$$

for all $k \geqq 1$. Denote by F_k the distribution of X_k and let $\{G_k, k \geqq 1\}$ be a sequence of distributions on B_k such that

$$F_k(A) \leqq G_k(A^{\rho_k}) + \sigma_k \quad \textit{for all } A \in \mathfrak{B}_k. \tag{1.6}$$

Here ρ_k *and* σ_k *are non-negative numbers and* $A^\varepsilon = \bigcup_{x\in A}\{y: m_k(x,y) < \varepsilon\}$. *Then without changing its distribution we can redefine the sequence* $\{X_k, k \geqq 1\}$ *on a richer probability space on which there exists a sequence* $\{Y_k, k \geqq 1\}$ *of independent random variables* Y_k *with distribution* G_k *such that for all* $k \geqq 1$

$$P\{m_k(X_k, Y_k) \geqq 2(\beta_k^{1/2} + \rho_k)\} \leqq 2(\beta_k^{1/2} + \sigma_k). \tag{1.7}$$

The proof of Theorem A can be found in Dehling and Philipp (1982). It is practically the same as the proof of similar approximation theorems of Berkes and Philipp (1979) and Philipp (1979).

In order to prove estimates for the Prohorov distance in the central limit theorem we need the following Theorem of Yurinskii (1977).

Theorem B. *Let* $X_1, X_2, \ldots, X_n$ *be independent* $\mathbb{R}^d$*-valued random variables with* $EX_i = 0$, $E\|X_i\|^3 < \infty$. *Denote the distribution of* $n^{-1/2}\sum_{i=1}^n X_i$ *by* μ_n *and let* ν_n *be the Gaussian measure with mean zero and same covariance as* μ_n. *Then:*

$$\pi(\mu_n, \nu_n) \leqq c d^{1/4} \rho_3^{1/4} n^{-1/8}(1 + |\log(\rho_3 n^{-1/2} d^{-1})|^{1/2}) \tag{1.8}$$

where $\rho_3 = n^{-1}\sum_{i=1}^n E\|X_i\|^3$ *and* c *is an absolute constant.*

Our first theorem is a partial generalization of Theorem B to weakly stationary strong mixing sequences of random variables with values in a finite-dimensional Banach space B.

A sequence $\{X_\nu, \nu \geqq 1\}$ of B-valued random variables is called weakly stationary if

$$Ef(X_n)\,g(X_m) = Ef(X_{n+k})\,g(X_{m+k}) \quad \text{for all } n, m \geqq 1,\ k \geqq 0,\ f, g \in B^*. \tag{1.9}$$

Theorem 1. a) *Let* $\{\xi_\nu, \nu \geqq 1\}$ *be a weakly stationary strong mixing sequence of random variables with values in a d-dimensional Banach space* B, *centered at expectations and with* $(2+\delta)$*-th moments uniformly bounded by* $\rho_{2+\delta}$, *where* $0 < \delta \leqq \frac{2}{3}$. *Suppose that the mixing coefficients* $\alpha(k)$ *satisfy*

$$\alpha(k) \ll k^{-(1+\varepsilon)(1+2/\delta)} \quad \textit{for some } 0 < \varepsilon \leqq 1. \tag{1.10}$$

Then the two series defining the covariance function T *of the sequence* $\{\xi_\nu, \nu \geqq 1\}$ *which is defined by:*

$$T(f,g) = Ef(\xi_1)\,g(\xi_1) + \sum_{k\geqq 2} Ef(\xi_1)\,g(\xi_k) + \sum_{k \geqq 2} Ef(\xi_k)\,g(\xi_1) \quad f, g \in B^* \tag{1.11}$$

converge absolutely. Moreover, there exist constants $\lambda > 0$ *and* C, *depending only on* ε, δ *and the constant implied by* $\ll$ *in* (1.10), *such that*

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^n \xi_j\right);\ N(0,T)\right) \leqq C n^{-\lambda} d^3 (1 + \rho_{2+\delta}^{1/3}). \tag{1.12}$$

In particular, if B is an inner product space, (1.12) *can be replaced by:*

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^{n}\xi_j\right);\ N(0,T)\right)\leqq Cn^{-\lambda}d^{3/2}(1+\rho_{2+\delta}^{1/3}). \tag{1.13}$$

b) *If the sequence* $\{\xi_\nu, \nu\geqq 1\}$ *is* ϕ*-mixing with rate*

$$\phi(k)\ll k^{-2(1+\varepsilon)} \quad \text{for some } 0<\varepsilon<1. \tag{1.14}$$

then the conclusions above remain valid. As a matter of fact we can replace (1.12) *and* (1.13) *by*

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^{n}\xi_j\right);\ N(0,T)\right)\leqq Cn^{-\lambda_1}d^{4/3}(1+\rho_{2+\delta}^{1/3}) \tag{1.15}$$

and

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^{n}\xi_j\right);\ N(0,T)\right)\leqq Cn^{-\lambda_1}d^{1/3}(1+\rho_{2+\delta}^{1/3}) \tag{1.16}$$

respectively.

Remarks. As a matter of fact the calculations yield

$$\lambda=\frac{\delta\varepsilon}{200(2+\varepsilon)} \tag{1.17}$$

and

$$\lambda_1=(\varepsilon\delta)/144. \tag{1.18}$$

Theorem 2. *Let* (X_j) *be a weakly stationary strong mixing sequence of random variables with values in the separable Banach space B such that* $EX_j=0$, $\sup_j E\|X_j\|^{2+\delta}\leqq\rho_{2+\delta}<\infty$ *for some* $0<\delta\leqq\frac{2}{3}$ *and suppose that the mixing coefficients satisfy* (1.10).

Let P_N *be a sequence of bounded operators on B with N-dimensional range. Assume that:*

$$E\left\|n^{-\frac{1}{2}}\sum_{j=a+1}^{a+n}(X_j-P_NX_j)\right\|^2\ll N^{-s} \quad \text{for some } s>0 \text{ uniformly in } a\geqq 0,\ n\geqq 1, \tag{1.19}$$

and

$$\|P_N\|=\sup_{\|x\|=1}\|P_Nx\|\ll N^r \quad \text{for some } r\geqq 0. \tag{1.20}$$

Then the two series defining the covariance function of T, which is defined as in (1.11), *converge absolutely. Moreover* $N(0,T)$ *exists and we have:*

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{i=1}^{n}X_i\right),\ N(0,T)\right)\ll(1+\rho_{2+\delta}^{1/3})n^{-\kappa} \tag{1.21}$$

where $\kappa=\dfrac{s\lambda}{9+3r+s}$ *and* λ *as in* (1.17).

The constant implied by $\ll$ *in* (1.21) *depends only on* ε, δ *and the constants implied by* $\ll$ *in* (1.10), (1.19), (1.20).

Moreover, the conslusions remain valid for ϕ-mixing sequences satisfying (1.14).

Using Theorem A we can easily obtain from Theorem 2 the following almost sure invariance principle:

Theorem 3. *Let $\{X_j, j \geqq 1\}$ be a weakly stationary sequence of random variables with values in the separable Banach space B such that $EX_j = 0$, $\sup_j E\|X_j\|^{2+\delta} \leqq \rho_{2+\delta} < \infty$ for some $0 < \delta \leqq \frac{2}{3}$. Suppose that $\{X_j; j \geqq 1\}$ is either ϕ-mixing with mixing rate* (1.14) *or absolutely regular with mixing rate*

$$\beta(k) \ll k^{-(1+\varepsilon)(1+2/\delta)} \quad \text{for some } \varepsilon > 0. \tag{1.22}$$

Let P_N be as in Theorem 2. Then we can redefine the sequence $\{X_j, j \geqq 1\}$ on a new probability space together with Brownian motion $X(t)$ with covariance structure T such that

$$\|\sum_{j \leqq t} X_j - X(t)\| \ll t^{1/2-\gamma} \quad \text{for some } \gamma > 0.$$

Note that, although we have an estimate for the Prohorov-distance in the central limit theorem for strong mixing sequences, we are not able to prove an invariance principle as in Theorem 3 for this case. This is due to the fact that Theorem A does not hold for strong mixing sequences. But a "good" estimate for the speed of convergence in the central limit theorem together with a maximal inequality always imply the law of the iterated logarithm. This on the other hand, implies an almost sure invariance principle with error term $o(t \log\log t)^{1/2}$, as was remarked in Dehling and Philipp (1982).

Theorem 4. *Let $\{X_j, j \geqq 1\}$ be a weakly stationary strong mixing sequence of random variables with values in the separable Banach space B such that $EX_j = 0$, $\sup_j E\|X_j\|^{2+\delta} < \infty$ for some $0 < \delta \leqq \frac{2}{3}$ and suppose that the mixing coefficients $\alpha(k)$ satisfy*

$$\alpha(k) \ll k^{-(2+\varepsilon)(1+2/\delta)} \quad \text{for some } \varepsilon > 0. \tag{1.23}$$

Let P_N be a sequence of bounded operators on B with N-dimensional range and assume that (1.20) *holds and that*

$$E\|n^{-1/2} \sum_{j \leqq n} (X_j - P_N X_j)\|^2 \ll (\log N)^{-3\alpha} \quad \text{for some } \alpha > 1. \tag{1.24}$$

Then we can redefine the sequence $\{X_j, j \geqq 1\}$ on a new probability space together with Brownian motion $X(t)$ with covariance structure T such that

$$\|\sum_{j \leqq t} X_j - X(t)\| = o(t \log\log t)^{1/2}.$$

Remark. For Hilbert space valued random variables this theorem was proved in [6] under much less restrictive assumptions, namely without (1.24) and with (1.23) replaced by (1.10).

If the random variables have values in the separable Hilbert space H, then Theorem 2 and 3 assume a simpler form.

Let $\{X_j, j \geqq 1\}$ be a strictly stationary sequence of H-valued random variables centered at expectations and with finite $(2+\delta)$-th moments. Suppose that $\{X_j, j \geqq 1\}$ is either ϕ-mixing with mixing rate (1.14) or absolutely regular with mixing rate (1.22). As a consequence of Theorem 8 and 9 in these cases the Gaussian measure $N(0, T)$ exists, where T is defined as in (1.11).

It is well known that T can be regarded as an element of $B(H, H)$, the space of bounded linear operators from H into itself. Moreover T is compact, selfadjoint and has finite trace. Hence there exists an orthonormal basis $(\tilde{e}_i)$ of eigenvectors of T with corresponding eigenvalues $(\tilde{\lambda}_i)$ for H. Denote the system consisting of those $\tilde{e}_i$ with $\tilde{\lambda}_i \neq 0$ by (e_i) and let (λ_i) be the corresponding eigenvalues. Then every $x \in H$ can be written as:

$$x = \sum_i (x, e_i) e_i + L(x)$$

where L is the orthogonal projection from H onto the orthogonal complement of the closed linear span of the e_i's.

The following corollary is an improvement of Corollary 3 of Kuelbs and Philipp (1980).

Corollary 1. *Let $\{X_j, j \geqq 1\}$ be as above. Moreover suppose that with the e_i's defined as above we have in the ϕ-mixing case*

$$\sum_{i \geqq N} E(X_1, e_i)^2 \ll N^{-s} \quad \textit{for some } s > 0 \tag{1.25}$$

and in the absolutely regular case

$$(E(\sum_{i \geqq N} (X_1, e_i)^2)^{1+\delta/2})^{2/(2+\delta)} \ll N^{-s} \quad \textit{for some } s > 0. \tag{1.25'}$$

Then (1.21) *holds with* $\kappa = \frac{2s}{2+2s}$. *Moreover the conclusion of Theorem* 3 *remains valid.*

Of course the statements of Theorem 2 and 3 as well as Corollary 1 contain central limit theorems for weakly stationary sequences. For strictly stationary sequences $\{X_j, j \geqq 1\}$ with values in a separable Banach space we can prove central limit theorems under much weaker conditions. If the sequence $\{X_j, j \geqq 1\}$ is ϕ-mixing, we need only finite second moments and that the mixing coefficients $\phi(k)$ satisfy:

$$\sum_{k \geqq 1} \phi^{1/2}(k) < \infty. \tag{1.26}$$

If the sequence is strong mixing, we need still finite $(2+\delta)$-th moments, but (1.10) can be replaced by

$$\sum_{k \geqq 1} (\alpha(k))^{\delta/(2+\delta)} < \infty. \tag{1.27}$$

In the special case where B is a separable Hilbert space our theorems generalize theorems of Ibragimov (1962).

Theorems 2, 3 and 4 will be applied to an example that was already treated before by Kaufman and Philipp (1978) and Kuelbs and Philipp (1980).

Let $\alpha > 1/2$ and define

$$\Lambda_\alpha = \left\{ f \in C[0,1] : |f(x) - f(y)| \leqq |x-y|^\alpha \quad x, y \in [0,1], \quad f(0) = f(1), \ \int_0^1 f\, dx = 0 \right\}. \tag{1.28}$$

By Arzela-Ascoli's theorem Λ_α is a compact subset of $C[0,1]$ and hence $C(\Lambda_\alpha)$, the class of all continuous real-valued functions on Λ_α is a separable Banach-space with the usual sup-norm.

Now let $\{\xi_j, j \geqq 1\}$ be a strictly stationary sequence of random variables uniformly distributed over $[0,1]$. Moreover, assume that $\{\xi_i, i \geqq 1\}$ is either ϕ-mixing satisfying (1.14) or absolutely regular satisfying (1.22) or strong mixing satisfying (1.10).

We define now random variables X_j with values in $C(\Lambda_\alpha)$ in the same way as Kuelbs and Philipp (1980).

$$X_j(\omega, f) = f(\xi_j(\omega)). \tag{1.29}$$

The following result is a partial improvement of Theorem 6 of Kuelbs and Philipp (1980).

Theorem 5. *Let X_i be defined as in* (1.29). *Then there exists a Gaussian measure μ on $C(\Lambda_\alpha)$ such that:*

$$\pi\left(\mathscr{L}\left(n^{-1/2} \sum_{j=1}^{n} X_j\right); \mu\right) \ll n^{-\kappa} \quad \textit{for some } \kappa > 0.$$

a) *If $\{\xi_j, j \geqq 1\}$ is ϕ-mixing or absolutely regular, we can redefine $\{X_j, j \geqq 1\}$ on a new probability space together with $C(\Lambda_\alpha)$-valued Brownian motion $X(t)$ such that*

$$\left\| \sum_{j \leqq t} X_j - X(t) \right\| \ll t^{\frac{1}{2} - \lambda} \quad \textit{for some } \lambda > 0.$$

b) *If $\{\xi_j, j \geqq 1\}$ is strong mixing satisfying* (1.23), *we can redefine $\{X_j, j \geqq 1\}$ on a new probability space together with $C(\Lambda_\alpha)$-valued Brownian motion $X(t)$ such that*

$$\left\| \sum_{j \leqq t} X_j - X(t) \right\| = o((t \log\log t)^{1/2}).$$

As a second example we shall consider $C(S)$-valued random variables. Let (S, τ) be a compact metric space and let $C(S)$ be the set of all continuous real-valued functions on S. $C(S)$ equipped with the sup-norm is a separable Banach space.

Since S is compact, there exists for every $\varepsilon > 0$ a finite number of ε-balls that cover S. Denote the minimal number by $N_\tau(S, \varepsilon)$. $N_\tau(S, \varepsilon)$ is a non-increasing function and hence there exists a (not necessarily unique) inverse function $g(x)$, say.

Let $\{X_j, j \geqq 1\} = \{X_j(s), j \geqq 1\}$ be a weakly stationary sequence of $C(S)$-valued random variables centered at expectations and with $(2+\delta)$-th moments uniformly bounded by $\rho_{2+\delta}$. Again suppose that $\{X_j, j \geqq 1\}$ is ϕ-mixing or absolutely regular with the usual mixing rates.

Let $R_n(s) = \sum_{j \leqq n} X_j(s)$ and suppose

$$\sup_{n \geqq 1} E \sup \{n^{-1} |R_n(s) - R_n(s')|^2 : \tau(s, s') \leqq \varepsilon, \quad s, s' \in S\} \leqq u(\varepsilon) \downarrow 0 \tag{1.30}$$

as $\varepsilon \to 0$. Then the sequence $\{n^{-1/2} R_n; n \geqq 1\}$ is tight and hence with the usual arguments one can show that $\{X_j, j \geqq 1\}$ satisfies the central limit theorem. If, moreover, $u \circ g(x)$ goes to zero fast enough as $x \to \infty$ we can even improve this and get the following result:

Theorem 6. *With the notation as above assume*

$$u(g(n)) \ll n^{-s} \quad \text{as} \quad n \to \infty \quad \text{for some } s > 0. \tag{1.31}$$

Then (1.21) *holds with* $\kappa = \dfrac{2s\lambda}{9+2s}$ *and the conclusion of Theorem* 3 *remains valid.*

Remark. In a forthcoming paper Marcus and Philipp (1982) prove similar almost sure invariance principles, but only for independent random variables.

As a by-product of the proof of Theorem 1 we obtain the following result which is useful in applications:

Theorem 7. *Let μ and ν be two Gaussian measures on $\mathbb{R}^d$ with mean* 0 *and covariance operators T and S. Then the following estimation for their Prohorov-distance holds:*

$$\pi(\mu, \nu) \leqq C \|T-S\|_1^{1/3} d^{1/6} (1 + \log(\|T-S\|_1^{-1} d)|^{1/2}) \tag{1.32}$$

where C is an absolute constant.

Remark. In (1.32) $\|A\|_1$ denotes the trace class norm of A, which can be defined for compact operators on a separable Hilbert-space H. The definition of $\| \ \|_1$ goes as follows: (for the details see Kuo (1975.)

A can be written in the form $A = UT$, where T is a positive compact operator on H and U is an isometry mapping the range of T into H. Then

$$\|A\|_1 = \sum_{n=1}^{\infty} \lambda_n,$$

where the λ_n's are the eigenvalues of T.

If A is also self-adjoint, there is a simpler representation of $\|A\|_1$.

Lemma 1.1. *For a compact and self-adjoint operator A on H we have:*

$$\|A\|_1 = \sup_{(e_n)} \sum_{n=1}^{\infty} |\langle A e_n, e_n \rangle|, \tag{1.33}$$

where the supremum is taken over all orthonormal bases (e_n) of H.

Proof. i) If A is in addition positive definite, (1.33) trivially holds: In this case we have for every complete orthonormal system (e_n):

$$\sum_{n=1}^{\infty} |\langle A e_n, e_n\rangle| = \sum_{n=1}^{\infty} \langle A e_n, e_n\rangle = \sum_{n=1}^{\infty} \lambda_n = \|A\|_1.$$

ii) In general, we have the following spectral representation of A:

$$A x = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n,$$

where the e_n's are the eigenvectors of A and the λ_n's are the corresponding eigenvalues. We define:

$$A^+(x) = \sum_{n=1}^{\infty} \lambda_n^+ \langle x, e_n\rangle e_n \qquad A^-(x) = \sum_{n=1}^{\infty} \lambda_n^- \langle x, e_n\rangle e_n,$$

where $\lambda_n^+ = \max(\lambda_n, 0)$ and $\lambda_n^- = \max(-\lambda_n, 0)$.

One can easily see that $\|A\|_1 = \sum_{n=1}^{\infty} |\lambda_n| = \|A^+\|_1 + \|A^-\|_1$ and hence:

$$\begin{aligned}\|A\|_1 = \sum_{n=1}^{\infty} |\lambda_n| &\leqq \sup_{(e_n)} \sum_{n=1}^{\infty} |\langle A e_n, e_n\rangle| \\ &\leqq \sup_{(e_n)} \sum_{n=1}^{\infty} \langle A^+ e_n, e_n\rangle + \sup_{(e_n)} \sum_{n=1}^{\infty} \langle A^- e_n, e_n\rangle \\ &= \|A^+\|_1 + \|A^-\|_1 = \|A\|_1\end{aligned}$$

which proves (1.33). □

Let us finally briefly indicate the contents of the various sections. Section 2 gives estimates for the Prohorov-distance of two Gaussian measures with different covariances. Section 3 contains lemmas on mixing random variables. Most of them are known, but we list them for the sake of completeness. Section 4 gives central limit theorems, weak and distribution type invariance principles for mixing sequences. In Sect. 5 we prove Theorem 1, part (b), by combining Theorem A and B. Since its proof does not work for strong mixing sequences, we need a different proof for part (a) of Theorem 1, which we give in Sect. 6. Section 7 contains the proof of Theorem 2, Sect. 8 the proof of Theorem 3 and Corollary 1. Theorem 4 is proved in Sect. 9 and in Sect. 10 and 11 we treat the examples given in Theorem 5 and Theorem 6.

2. Estimates for the Prohorov-distance of Gaussian Measures

Later in this paper we shall often face the problem that we have to estimate the Prohorov-distance of two mean zero Gaussian measures with known covariances. A first almost trivial, but nevertheless sometimes useful result in this direction is

Lemma 2.1. *Let X and Y be $\mathbb{R}^d$-valued square integrable random variables with mean zero. Then we have the following estimation for the Prohorov-distance of the corresponding Gaussian measures:*

$$\pi(N(0, \operatorname{cov} X); N(0, \operatorname{cov} Y)) \leqq (E\|X-Y\|^2)^{1/3}. \tag{2.1}$$

Proof. We consider the random variable (X, Y) in $\mathbb{R}^{2d}$. (X, Y) is square integrable and hence there exists a Gaussian measure μ on $\mathbb{R}^{2d}$ with mean zero and same covariance. Let Z be an $\mathbb{R}^{2d}$-valued random variable with distribution μ and denote by pr_1 and pr_2 the projections of $\mathbb{R}^{2d}$ on the first d and the last d coordinates.

Then the r.v.'s $\mathrm{pr}_1 Z$, $\mathrm{pr}_2 Z$, $\mathrm{pr}_1 Z$-$\mathrm{pr}_2 Z$ have alle Gaussian distributions with the same covariance as X, Y, $X-Y$.

By Čebyšev's inequality and the fact that $E\|X\|^2 = \operatorname{trace}(\operatorname{cov} X)$ for any mean-zero random variable X, we get:

$$\begin{aligned}\pi(N(0, \operatorname{cov} X); N(0, \operatorname{cov} Y)) &\leqq (E\|\mathrm{pr}_1 Z - \mathrm{pr}_2 Z\|^2)^{1/3} \\ &= (\operatorname{trace}(\operatorname{cov}(\mathrm{pr}_1 Z - \mathrm{pr}_2 Z)))^{1/3} \\ &= (E\|X-Y\|^2)^{1/3}. \quad \square\end{aligned}$$

Unfortunately, in most problems we know only the covariance matrix of the Gaussian measures, but nothing about the L^2-distance of related random variables. In order to obtain estimates for the Prohorov-distance in this case, we shall first prove a general result about Gaussian measures in Banach spaces.

For this we recall the notion of differentiability of functions with domain and range in a Banach space.

For two Banach spaces E, F we define:

$$L(E, F) = \{\phi : E \to F \mid \phi \text{ is bounded and linear}\}.$$

A function f: $E \to F$ is said to be differentiable at the point $x_0 \in E$, if there exists a $\phi \in L(E, F)$, such that:

$$\lim_{h \to 0} \frac{\|f(x_0+h) - f(x_0) - \phi(h)\|}{\|h\|} = 0.$$

In this case we define the derivative $D_{x_0} f$ of f at x_0 to be ϕ. In the same way we can define higher derivatives of f, which we denote by $D_x^n f$. For fixed $x \in E$ the n-th derivative $D_x^n f$ is an element of $L(E, \ldots, E; F)$, the space of all continuous n-linear functions from $E \times \ldots \times E$ into F. (See also Cartan (1971).)

For n-times differentiable functions $f: E \to \mathbb{R}$ an analogue of Taylor's theorem is valid.

Lemma 2.2. *Let $f: E \to \mathbb{R}$ be n times differentiable and let $x, h \in E$. Then*

$$f(x+h) = f(x) + D_x f(h) + \frac{1}{2!} D_x^2 f(h, h) + \ldots + \frac{1}{(n-1)!} D_x^{n-1} f(h, \ldots, h) + R_n(x, h, f) \tag{2.2}$$

where

$$R_n(x,h,f) \leqq \frac{1}{n!} \sup_x \|D_x^n f\| \, \|h\|^n.$$

For the proof of this lemma see Cartan (1971), Theorem 5.6.2.

We can now prove a result about the difference of the integrals of certain functions with respect to two Gaussian measures, which will be a key in the proof of Theorem 7.

Proposition 2.1. *Let E be a separable Banach space and let $f: E \to \mathbb{R}$ be a three times differentiable function with bounded third derivative. If μ and ν are Gaussian measures on E with mean zero, we have:*

$$\left|\int_E f\,d\mu - \int_E f\,d\nu\right| \leqq \tfrac{1}{2} \sup_{x\in E} \left|\int_E D_x^2 f(y,y)\,d(\mu-\nu)(y)\right|. \tag{2.3}$$

Proof. We shall prove this theorem by using the so-called operator method, due to Trotter (1959). Let μ be a probability measure on E; we define an operator W_μ on C_b^0, the class of all bounded real-valued equicontinuous functions on E, in the following way:

$$W_\mu f(x) = \int_E f(x+y)\,d\mu(y).$$

If $\mu, \mu_1, \ldots, \mu_n$, ν, $\nu_1, \ldots, \nu_n$ are probability measures on E, the following facts are easily proved:

$$\begin{aligned} &\text{i) } W_\mu f \in C_b^0; \quad \|W_\mu\| \leqq 1,\\ &\text{ii) } W_{\mu*\nu} = W_\mu \circ W_\nu,\\ &\text{iii) } \|W_{\mu_1*\cdots*\mu_n} f - W_{\nu_1*\cdots*\nu_n} f\| \leqq \sum_{i=1}^n \|W_{\mu_i} f - W_{\nu_i} f\|. \end{aligned} \tag{2.4}$$

Now let S and T be the covariance operators for the Gaussian measures μ and ν. We define:

$$\mu_n = N(0, n^{-1} S), \qquad \nu_n = N(0, n^{-1} T). \tag{2.5}$$

Hence:

$$\mu = \mu_n * \ldots * \mu_n \qquad \nu = \nu_n * \ldots * \nu_n. \tag{2.6}$$

Using (2.4) and (2.6) we conclude:

$$\begin{aligned} \left|\int f(x)\,d\mu(x) - \int f(x)\,d\nu(x)\right| &= |W_\mu f(0) - W_\nu f(0)|\\ &\leqq \|W_\mu f - W_\nu f\| \leqq n \|W_{\mu_n} f - W_{\nu_n} f\|. \end{aligned} \tag{2.7}$$

By definition we have

$$\|W_{\mu_n} f - W_{\nu_n} f\| = \sup_x \left|\int f(x+y)\,d(\mu_n - \nu_n)(y)\right|. \tag{2.8}$$

We fix $x \in E$, and using the Taylor-expansion of f about x and the linearity of D_x we get:

$$\begin{aligned} \int f(x+y)\,d(\mu_n-\nu_n)(y) &= \int f(x) + D_x f(y) + \tfrac{1}{2} D_x^2 f(y,y) + R_3(x,y,f)\,d(\mu_n-\nu_n)(y)\\ &= \tfrac{1}{2} \int D_x^2 f(y,y)\,d(\mu_n-\nu_n)(y) + \int R_3(x,y,f)\,d(\mu_n-\nu_n)(y). \end{aligned} \tag{2.9}$$

By (2.5) and bilinearity of $D_x^2 f$ we have then:

$$\int D_x^2 f(y,y)\, d(\mu_n - \nu_n)(y) = \frac{1}{n}\int D_x^2 f(y,y)\, d(\mu-\nu)(y). \tag{2.10}$$

From (2.2) we get:

$$\begin{aligned}|\int R_3(x,y,f)\, d(\mu_n-\nu_n)(y)| &\leqq \tfrac{1}{6}\sup_x \|D_x^3 f\| \int \|y\|^3\, d(\mu_n+\nu_n)(y)\\ &= \tfrac{1}{6}\sup_x \|D_x^3 f\|\, n^{-3/2}\int \|y\|^3\, d(\mu+\nu)(y)\\ &= C n^{-3/2}\end{aligned} \tag{2.11}$$

where C is independent of n. Combining (2.7), (2.8), (2.9), (2.10) and (2.11) we obtain:

$$|\int f(x)\, d\mu(x) - \int f(x)\, d\nu(x)| \leqq \tfrac{1}{2}\sup_x |\int D_x^2 f(y,y)\, d(\mu-\nu)(y)| + C n^{-1/2}.$$

Passing to the limit finally yields (2.3). □

The usefulness of Proposition 2.1 lies in the fact that, at least in Euclidean spaces, the right side of 2.3 can be estimated in terms of the covariance operators of μ and ν:

In what follows let E be the d-dim Euclidean space $\mathbb{R}^d$. By Schwartz's theorem we know that $D_x^2 f$ is symmetric, hence applying the principal axis theorem gives us the existence of an orthonormal basis $e_1^x, \ldots, e_d^x$ for $\mathbb{R}^d$ such that:

$$D_x^2 f(y,y) = \sum_{i=1}^{d} \lambda_i^x (y, e_i^x)^2$$

where λ_i^x are the eigenvalues for $D_x^2 f$. So if T and S are the covariance-operators for the Gaussian measures μ and ν, we get using (2.3) and (1.33):

$$\begin{aligned}|\int f\, d\mu - \int f\, d\nu| &\leqq \tfrac{1}{2}\sup_x \sum_{i=1}^{d} |\lambda_i^x|\, |((T-S)e_i^x; e_i^x)|\\ &\leqq \tfrac{1}{2}\sup_x \|D_x^2 f\|\, \|T-S\|_1.\end{aligned}$$

We summarize this result and get:

Corollary 2.1. *If in Proposition 2.1 E is replaced by $\mathbb{R}^d$ and T and S are the covariance-operators of μ and ν, then we have:*

$$|\int_{\mathbb{R}^d} f\, d\mu - \int_{\mathbb{R}^d} f\, d\nu| \leqq \tfrac{1}{2}\sup_{x\in\mathbb{R}^d} \|D_x^2 f\|\, \|T-S\|_1. \tag{2.12}$$

In order to obtain an estimate for the Prohorov-distance of μ and ν we will apply Corollary 2.1. to the smooth "indicators" of sets, as defined by Yurinskii (1977).

Lemma 2.3. *Let $A \subset \mathbb{R}^d$ be a closed set, then for any $\varepsilon \in (0,1)$ there exists a three times continuously differentiable function $\phi(x)$, such that:*

$$\begin{aligned} &0 \leqq \phi(x) \leqq 1 \\ &\phi(x) \geqq 1-2\varepsilon \quad \text{for } x \in A \\ &\phi(x) \leqq 2\varepsilon \quad \text{for } x \in \mathbb{R}^d \setminus A^{\varepsilon_1} \text{ where } \varepsilon_1 = 4\varepsilon \left(\log \frac{1}{\varepsilon}\right)^{1/2} d^{1/2} + 4\varepsilon d^{1/2}, \end{aligned} \tag{2.13}$$

$$\begin{aligned} &\|D_x^3 \phi\| \leqq 2\varepsilon^{-3} d^{-1/2} \\ &\|D_x^2 \phi\| \leqq \varepsilon^{-2} d^{-1/2} \\ &\|D_x \phi\| \leqq \varepsilon^{-1} d^{-1/2}. \end{aligned} \tag{2.14}$$

Proof. All the statements of the lemma except the last two inequalities are proved in Yurinskii (1977), Lemma 4, thus we only have to estimate $\|D_x^2 \phi\|$ and $\|D_x \phi\|$.

Recall from Yurinskii (1977)

$$\phi(x) = \int_{\mathbb{R}^d} \hat{\phi}(x+z)\, g(z)\, dz$$

where $g(z) = (2\pi\varepsilon^2)^{-d/2} \exp\left(-\frac{(z,z)}{2\varepsilon^2}\right)$ and where $\hat{\phi}$ is a.e. differentiable with $\|D_x \hat{\phi}\| \leqq d^{-1/2} \varepsilon^{-1}$.

As is proved in Yurinskii (1977), we have

$$D_x \phi(u) = \int D_{z+x} \hat{\phi}(u)\, g(z)\, dz, \tag{2.15}$$

$$D_x^2 \phi(u,v) = -\int D_z \hat{\phi}(u)\, D_{z-x} g(v)\, dz. \tag{2.16}$$

Using (2.15) we get the last inequality in (2.14). Using (2.16) and the fact that

$$D_z g = -\frac{1}{\varepsilon^2} g(z) \cdot z$$

we get

$$\begin{aligned} |D_x^2 \phi(u,v)| &\leqq \int |D_z \hat{\phi}(u)| \left| \frac{1}{\varepsilon^2}(z-x, v)\, g(z-x) \right| dz \\ &\leqq \|u\|\, d^{-1/2} \varepsilon^{-1} \varepsilon^{-2} \int |(z,v)|\, g(z)\, dz \\ &\leqq \|u\|\, d^{-1/2} \varepsilon^{-3} \left(\int (z,v)^2 g(z)\, dz\right)^{1/2} \\ &= \|u\|\, d^{-1/2} \varepsilon^{-3} (\varepsilon^2 \|v\|^2)^{1/2} \\ &= \|u\|\, \|v\|\, d^{-1/2} \varepsilon^{-2}. \quad \square \end{aligned}$$

The following almost trivial lemma enables us to prove Theorem 7.

Lemma 2.4. *Let μ, ν be two Borel measures on $\mathbb{R}^d$, $A \subset \mathbb{R}^d$ a Borel set. If $\phi: \mathbb{R}^d \to [0,1]$ is a measurable function such that:*

$$\phi(x) \geqq 1 - 2\varepsilon \quad \text{for } x \in A,$$
$$\phi(x) \leqq 2\varepsilon \quad \text{for } x \in \mathbb{R}^d / A^{\varepsilon_1}$$

then $\mu(A) \leqq \nu(A^{\varepsilon_1}) + 4\varepsilon + \delta$
where $\delta := |\int \phi \, d(\mu - \nu)|$.

Proof of Theorem 7. Take $\varepsilon = \|T-S\|_1^{1/3} d^{-1/3}$ and observe that in case $\varepsilon \geqq 1$, (1.32) is trivially valid. Thus we can assume without loss generality $\varepsilon < 1$.

Let ϕ be as in Lemma 2.3. By (2.12) and (2.14)

$$|\int \phi \, d\mu - \int \phi \, d\nu| \leqq \tfrac{1}{2} \varepsilon^{-2} d^{-1/2} \|T-S\|_1 = \tfrac{1}{2} d^{1/6} \|T-S\|_1^{1/3}.$$

Then we get by Lemma 2.4:

$$\begin{aligned}\pi(\mu,\nu) &\leqq 4\|T-S\|_1^{1/3} d^{-1/3} d^{1/2} (\log(\|T-S\|_1^{-1/3} d^{1/3}))^{1/2} + 4\|T-S\|_1^{1/3} d^{1/6} \\ &\quad + 4\|T-S\|_1^{1/3} d^{-1/3} + \tfrac{1}{2} d^{1/6} \|T-S\|_1^{1/3} \\ &\leqq C\|T-S\|_1^{1/3} d^{1/6} \left(1 + \left|\log \frac{d}{\|T-S\|_1}\right|^{1/2}\right).\end{aligned}$$

3. Some Lemmas on Mixing and Absolutely Regular Random Variables

Lemma 3.1. *Let $\mathfrak{F}$ and $\mathfrak{G}$ be two σ-fields. Define*

$$\alpha(\mathfrak{F}, \mathfrak{G}) = \sup |P(A \cap B) - P(A) P(B)|$$

the supremum being extended over all $A \in \mathfrak{F}$ and $B \in \mathfrak{G}$. Let ξ and η be random variables with values in a separable Hilbert space H measurable $\mathfrak{F}$ and $\mathfrak{G}$ respectively. If ξ and η are essentially bounded then

$$|E(\xi,\eta) - (E\xi, E\eta)| \leqq 10\alpha(\mathfrak{F},\mathfrak{G}) \|\xi\|_\infty \|\eta\|_\infty. \tag{3.1}$$

Here $\|\cdot\|_\infty$ denotes the essential supremum with respect to H. Moreover, let $r,s,t > 1$ with $r^{-1} + s^{-1} + t^{-1} = 1$. If ξ and η have finite r-th and s-th moments respectively then

$$|E(\xi,\eta) - (E\xi, E\eta)| \leqq 15\alpha^{1/t}(\mathfrak{F},\mathfrak{G}) \|\xi\|_r \|\eta\|_s. \tag{3.2}$$

For real-valued random variables (3.1) is due to Volkonskii and Rozanov (1959) with 10 replaced by 4 and (3.2) is due to Davydov (1970). The proof of Lemma 3.1 in the general case was given by Dehling and Philipp (1982).

Lemma 3.2. *Let $\{\xi_\nu, \nu \geqq 1\}$ be a weakly stationary sequence of H-valued random variables centered at expectations and with $(2+\delta)$-th moments uniformly bounded where $0 < \delta \leqq 1$. If $\{\xi_\nu, \nu \geqq 1\}$ satisfies a strong mixing condition with mixing rate (1.27), then*

$$E\left\|\sum_{\nu=a+1}^{a+n} \xi_\nu\right\|^2 = \sigma^2 n + o(n) \tag{3.3}$$

and in fact

$$E\left\|\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^2 \leqq n(1+\sum_{j\geqq 1}(\alpha(j))\delta/(2+\delta)\sup_{\nu\geqq 1}\|\xi_\nu\|_{2+\delta}^2 \tag{3.4}$$

where

$$\sigma^2 = E\|\xi_1\|^2 + 2\sum_{\nu\geqq 2}E(\xi_1,\xi_\nu).$$

For a proof of this see e.g. the proof of Lemma 2.3 in [14].

Remark. (3.4) holds without the assumption of weak stationarity.

Lemma 3.3. *Let $\{\xi_\nu, \nu\geqq 1\}$ be a sequence of random variables with values in a separable Banach space satisfying a strong mixing condition* (1.1) *with mixing rate*

$$\alpha(k) \ll k^{-(1+\varepsilon)(1+2/\delta)} \quad \textit{for } 0<\varepsilon<1. \tag{3.5}$$

Suppose that their $(2+\delta)$-th moments are uniformly bounded by b, where $0<\delta\leqq 1$ and that for some $\sigma<\infty$

$$E\left\|\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^2 \leqq \sigma^2 n \quad \textit{for all } a\geqq 0,\ n\geqq 1. \tag{3.6}$$

Then for all $a\geqq 0$ and all $0\leqq\alpha\leqq\varepsilon\delta/8$

$$E\left\|\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^{2+\alpha} \ll n^{1+\alpha/2}(\sigma^{2+\alpha}+b^{(2+\alpha)(2+\delta)^{-1}})$$

where the constant implied by $\ll$ only depends on ε, δ and the constant implied by $\ll$ in (3.5).

In the real-valued case Lemma 3.3 is due to Sotres and Malay Ghosh (1977). But as was already remarked in Dehling and Philipp (1982) their proof still works for B-valued random variables.

For ϕ-mixing sequences the three lemmas stated above can be strengthened considerably.

Lemma 3.4. *Let $\mathfrak{F}$ and $\mathfrak{G}$ be two σ-fields and define*

$$\phi(\mathfrak{F},\mathfrak{G}) = \sup|P(B|A)-P(B)|$$

the supremum being extended over all $A\in\mathfrak{F}$ and $B\in\mathfrak{G}$. Let ξ and η be random variables with values in a separable Hilbert space H measurable $\mathfrak{F}$ and $\mathfrak{G}$ respectively. If $p,q\geqq 1$ satisfy $p^{-1}+q^{-1}=1$ and if ξ and η have finite p-th and q-th moments respectively then

$$|E(\xi,\eta)-(E\xi,E\eta)| \leqq 2\phi^{1/p}(\mathfrak{F},\mathfrak{G})\|\xi\|_p\|\eta\|_q. \tag{3.7}$$

The proof of this lemma can be found for real-valued random variables in Billingsley (1968), p. 170, but his proof works also for H-valued random variables. The same is true for the following lemma.

Lemma 3.5. *Let $\{\xi_\nu, \nu \geqq 1\}$ be a weakly stationary sequence of H-valued random variables, centered at expectations and with finite second moments. Suppose that $\{\xi_\nu, \nu \geqq 1\}$ is ϕ-mixing and that the mixing coefficients $\phi(k)$ satisfy* (1.26). *Then*

$$E\|n^{-1/2}\sum_{\nu\leqq n}\xi_\nu\|^2=\sigma^2+o(1) \tag{3.8}$$

and

$$E\|n^{-1/2}\sum_{\nu\leqq n}\xi_\nu\|^2\leqq\left(1+\sum_{k=1}^{\infty}\phi^{1/2}(k)\right)E\|\xi_1\|^2. \tag{3.9}$$

Lemma 3.6. *Let B be a separable Banach space and $\{\xi_\nu, \nu \geqq 1\}$ be a ϕ-mixing sequence of B-valued random variables centered at expectations and with $(2+\delta)$-th moments uniformly bounded by $\rho_{2+\delta}$ where $0<\delta<1$. If for some $\sigma<\infty$*

$$E\left\|\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^2\leqq\sigma^2 n \qquad a\geqq 0,\ n\geqq 1 \tag{3.10}$$

holds, then we have

$$E\left\|\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^{2+\delta}\ll n^{1+\delta/2}(\sigma^{2+\delta}+\rho_{2+\delta}) \tag{3.11}$$

where the constant implied by $\ll$ depends only on ϕ and δ.

Remark. Note that the only requirement on ϕ is $\phi(k)\to 0$. Lemma 3.6 was proved by Kuelbs and Philipp (1980) under the additional restriction that $\phi(k)$ tends to zero at a certain rate.

Proof of Lemma 3.6. The proof follows completely the lines of Doob (1953), p. 225–227 and only a few minor changes are necessary. First note that without loss of generality we can assume $\sigma^{2+\delta}+\rho_{2+\delta}\leqq 1$. Then define

$$c_n=\sup_a\left\|E\sum_{\nu=a+1}^{a+n}\xi_\nu\right\|^{2+\delta}.$$

As in Doob we can show that $c_n\leqq bn^{1+\delta/2}$ for some constant b depending only on ϕ and δ. For the proof of (7.10) in Doob we have to use of course our Lemma 3.4. Finally note that because of our definition of c_n we do not need stationarity. □

Now that we have stated inequalities for strong and ϕ-mixing random variables, the reader may wonder whether there are similar inequalities for absolutely regular sequences. Unfortunately no special inequalities are known for these sequences and all we can do is to use the ones for strong mixing sequences. The only reason for the fact that our main results for absolutely regular sequences are worse than for ϕ-mixing sequences is the fact that the above lemmas yield worse estimates for the strong mixing than for the ϕ-mixing case.

Now let H be a separable Hilbert space and denote by $\mathcal{S}$ the set of all covariance operators of Gaussian measures on H. Let $\mathcal{S}$ be equipped with the metric d, defined by

$$d(S,T)=\|S-T\|_1$$

It is easy to see that $(\mathscr{S},d)$ is a complete metric space.

Lemma 3.7. a) *Let $\{\xi_i, i\geq 1\}$ be a weakly stationary ϕ-mixing sequence of H-valued random variables with $E\xi_i=0$, $E\|\xi_i\|^2<\infty$. Assume that the mixing coefficients satisfy* (1.14). *If the covariance-operator of $n^{-1/2}\sum_{i=1}^{n}\xi_i$ is denoted by T_n, then (T_n) converges in $(\mathscr{S},d)$ to a limit T, say. Moreover T is given by*

$$(Tx,y)=E(\xi_1,x)(\xi_1,y)+\sum_{k\geq 2}E(\xi_1,x)(\xi_k,y)+\sum_{k\geq 2}E(\xi_k,x)(\xi_1,y) \tag{3.12}$$

and the following estimate holds:

$$\|T_n-T\|_1 \ll n^{-\varepsilon}E\|\xi_1\|^2 \tag{3.13}$$

where the constant implied by $\ll$ depends only on the one implied by $\ll$ in (1.14) *and on ε.*

b) *Let $\{\xi_i, i\geq 1\}$ be a weakly stationary strong mixing sequence of $\mathbb{R}^d$-valued random variables centered at expectations and with $(2+\delta)$-th moments uniformly bounded by $\rho_{2+\delta}$. Assume that the mixing coefficients $\alpha(k)$ satisfy* (3.5). *Then the two series defining T as in* (3.12) *converge absolutely and the following estimate holds*

$$\|T_n-T\|_1 \ll d n^{-\varepsilon}\rho_{2+\delta}^{2(2+\delta)^{-1}}. \tag{3.14}$$

Here the constant implied by $\ll$ depends only on ε and the constant implied by $\ll$ in (3.5) *but not on d.*

Proof. a) First we show that T as defined in (3.12) belongs to $\mathscr{S}$. By Lemma 3.4 and weak stationarity we have that:

$$\begin{aligned}|(Tx,x)| &\leq E(\xi_1,x)^2+2\sum_{k\geq 2}E(\xi_1,x)^2(\phi(k-1))^{1/2}\\ &\ll E(\xi_1,x)^2\Big(1+2\sum_{k\geq 1}k^{-(1+\varepsilon)}\Big).\end{aligned}$$

Hence for each orthonormal basis $\{e_n, n\geq 1\}$:

$$\sum_{n=1}^{\infty}|(Te_n,e_n)| \ll E\|\xi_1\|^2\Big(1+2\sum_{k\geq 1}k^{-1-\varepsilon}\Big).$$

It remains to show (3.13). By weak stationarity we have, putting $S_n=n^{-1/2}\sum_{i\leq n}\xi_i$

$$\begin{aligned}(T_nx,y)&=E(S_n,x)(S_n,y)\\ &=E(\xi_1,x)(\xi_1,y)+\frac{1}{n}\sum_{k=2}^{n}(n-k+1)\,E(\xi_1,x)(\xi_k,y)\\ &\quad+\frac{1}{n}\sum_{k=2}^{n}(n-k+1)\,E(\xi_k,x)(\xi_1,y).\end{aligned}$$

Hence for any orthonormal basis (e_j) of H we have:

$$\begin{aligned}((T-T_n)e_j, e_j) &\leqq 2\left(\left|\sum_{k=n+1}^{\infty} E(\xi_1, e_j)(\xi_k, e_j)\right| + \left|\sum_{k=2}^{n} \frac{k-1}{n} E(\xi_1, e_j)(\xi_k, e_j)\right|\right)\\ &\leqq 2\left(\sum_{k=n+1}^{\infty} \phi^{1/2}(k-1) + \sum_{k=2}^{n} \frac{k-1}{n} \phi^{1/2}(k-1)\right) E(\xi_1, e_j)^2\\ &\ll \left(\sum_{k=n}^{\infty} k^{-1-\varepsilon} + \sum_{k=1}^{n-1} \frac{k}{n} k^{-1-\varepsilon}\right) E(\xi_1, e_j)^2\\ &\ll n^{-\varepsilon} E(\xi_1, e_j)^2.\end{aligned}$$

Summing over j and using (1.33) we get (3.13).

b) The absolute convergence of the series in (3.12) follows by the same kind of computations as in (a). Instead of Lemma 3.4 we just have to use Lemma 3.1 with $s=t=(2+\delta)^{-1}$ and $r=\delta(2+\delta)^{-1}$. In the same way we can show that

$$((T_n - T)e_j, e_j) \ll n^{-\varepsilon} \rho_{2+\delta}^{2(2+\delta)^{-1}}.$$

Summing over j then yields (3.14).

Lemma 3.8. *Let $\mathfrak{F}$ and $\mathfrak{G}$ be two σ-fields and define $\alpha=\alpha(\mathfrak{F}, \mathfrak{G})$ as in Lemma 3.1. If ξ is $\mathfrak{G}$-measurable and $E|\xi|^r < \infty$, then we have for $1<r<p$,*

$$E|E(\xi|\mathfrak{F}) - E\xi|^r \leqq 15\alpha^{1-\frac{r}{p}}(E|\xi|^p)^{\frac{r}{p}}. \tag{3.15}$$

Proof. For the proof we use a trick of Philipp and Stout (1975). Assume without loss of generality that $E\xi=0$. Then we get using (3.2)

$$\begin{aligned}E|E(\xi|\mathfrak{F})|^r &= E(E(\xi|\mathfrak{F}) \operatorname{sgn} E(\xi|\mathfrak{F})\, |E(\xi|\mathfrak{F})|^{r-1})\\ &= E(\xi \operatorname{sgn} E(\xi|\mathfrak{F})\, |E(\xi|\mathfrak{F})|^{r-1})\\ &\leqq 15\alpha^{1-\frac{r}{p}}(E|\xi|^p)^{1/p}(E|E(\xi|F)|^p)^{\frac{r-1}{p}}\\ &\leqq 15\alpha^{1-\frac{r}{p}}(E|\xi|^p)^{\frac{r}{p}}.\end{aligned}$$

Lemma 3.9. *Let $\{\xi_v, v\geqq 1\}$ be a strong mixing sequence of B-valued random variables, centered at expectations and with $(3+\delta)$-th moments uniformly bounded by $\rho_{3+\delta}$ for some $\delta>0$. Suppose that the mixing coefficients $\alpha(k)$ satisfy*

$$\alpha(k) \ll k^{-(1+\varepsilon)(3+9/\delta)} \quad \text{for some } \varepsilon>0. \tag{3.16}$$

If $\sigma^2 = \sup E\|n^{-1/2}\sum_{v\leqq n} \xi_v\|^2$, then we have

$$E\left\|\sum_{1}^{n} \xi_v\right\|^3 \leqq C n^{3/2}(\sigma^3 + \rho_{3+\delta}^{3/(3+\delta)}) \tag{3.17}$$

where C is a constant depending only on ε, δ and the constant implied by $\ll$ in (3.16).

Proof. We follow the lines of a similar proof in Kuelbs and Philipp ([14]). For given n define

$$k=[n^{\frac{1}{2}-\gamma}] \qquad m=[\tfrac{1}{2}n]-k$$

where γ is chosen in such a way that $(\frac{1}{2}-\gamma)(1+\varepsilon)>\frac{1}{2}$. Then $\alpha(2k)\ll n^{-\frac{1}{2}(3+9/\delta)}$. Define

$$R_a=\sum_{v=a+1}^{a+m}\xi_v, \qquad S_a=\sum_{v=a+m+2k+1}^{2m+2k}\xi_v.$$

Now put in Lemma 3.8. $r=\frac{3}{2}$, $p=\dfrac{3+\delta}{2}$ to get:

$$\begin{aligned} E|E(\|S_a\|^2|\mathfrak{M}_1^{a+m})-E\|S\|^2|^{3/2} &\leqq C n^{-\frac{1}{2}(3+9/\delta)(1-3/(3+\delta))}(E\|S_a\|^{3+\delta})^{3/3+\delta}\\ &\leqq C n^{-3/2} n^3\\ &\leqq C n^{3/2}. \end{aligned}$$

The rest of the proof follows now as in [14], p. 1016.

4. Some Central Limit Theorems

The following central limit theorems improve recent results of Kuelbs and Philipp (1980) and of Dehling and Philipp (1982).

Theorem 8. a) *Let $\{X_j, j\geqq 1\}$ be a stationary sequence of random variables with values in a separable Hilbert space H centered at expectations and with finite second moments. Suppose $\{X_j, j\geqq 1\}$ is ϕ-mixing and that the mixing coefficients $\phi(k)$ satisfy* (1.26). *Then the two series defining T, which is defined as in* (1.11), *converge absolutely and $n^{-1/2}\sum_{v\leqq n} X_v$ converges weakly to a Gaussian measure with covariance T.*

b) *Let $\{X_j, j\geqq 1\}$ be a stationary sequence of B-valued random variables with the same properties as in* a). *Moreover assume there exists for each ρ a measurable map $\Lambda_\rho: B\to B$ with finite-dimensional range, such that*

$$\begin{gathered} E\Lambda_\rho X_1=0, \qquad E\|\Lambda_\rho X_1\|^2<\infty\\ E\left\|n^{-1/2}\sum_{j=1}^n (X_j-\Lambda_\rho X_j)\right\|\leqq\rho. \end{gathered} \tag{4.1}$$

Then the conclusion of a) *hold.*

Theorem 9. *Theorem* 8 *continues to hold for strong mixing sequences with finite $(2+\delta)$-th moments if the mixing coefficients satisfy* (1.27) *and if* 4.1 *is replaced by*

$$\begin{gathered} E\Lambda_\rho X_1=0, \qquad E\|\Lambda_\rho X_1\|^{2+\delta}<\infty\\ E\left\|n^{-1/2}\sum_{j=1}^n (X_j-\Lambda_\rho X_j)\right\|\leqq\rho. \end{gathered} \tag{4.1'}$$

Proof of Theorem 8. b) For $\mathbb{R}$-valued random variables this is just Theorem 1.5 in Ibragimov (1962). Using the Cramér-Wold device we get the theorem if $B=\mathbb{R}^d$. Then we can proceed as in Proposition 4.2 of [14].

a) This can be proved as Corollary 2 in [14].

Proof of Theorem 9. This time for $\mathbb{R}$-valued random variables the result is just Theorem 1.7 in Ibragimov (1962) and the remaining part of b) follows as above. Part a) can again be proved as Corollary 2 in Kuelbs and Philipp (1980) if we replace their (4.33) by our (3.4). □

Although we have been unable to prove an almost sure invariance principle under the assumptions of Theorem 8 and 9, we still have a functional central limit theorem. We define a random element Z_n of $C_B[0,1]$, the space of all continuous functions f: $[0,1]\to B$, by:

$$Z_n(t,\omega)=\begin{cases} n^{-1/2}\sum_{j\leq k} X_j & \text{if } t=k/n \\ \text{linear in between.} \end{cases} \tag{4.2}$$

Theorem 10. *Let* $\{X_j, j\geqq 1\}$ *be a stationary sequence of random variables with values in B or H such that the assumptions in Theorem* 8 *hold. Then the sequence* $\{Z_n, n\geqq 1\}$ *defined in* (4.2) *converges in distribution to a B-valued Brownian motion on* $[0,1]$ *with covariance structure T.*

Proof. We apply Theorem 3.4. of Eberlein (1979). It is easily seen that the array $\{n^{-1/2}X_j,\ 1\leqq j\leqq n,\ n\geqq 1\}$ satisfies Eberleins conditions and this proves our theorem.

Theorem 11. a) *Let* $\{X_j, j\geqq 1\}$ *be a stationary sequence of random variables with values in a separable Hilbert space H centered at expectations and with finite* $(2+\delta)$*-th moments. Suppose* $\{X_j, j\geqq 1\}$ *satisfies a strong mixing condition* (1.1) *with rate of decay* (3.5). *Then the sequence* $\{Z_n, n\geqq 1\}$ *defined in* (4.2) *converges in distribution to a B-valued Brownian motion on* $[0,1]$ *with covariance structure T.*

b) *Let* $\{X_j, j\geqq 1\}$ *be a stationary sequence of B-valued random variables with the same properties as in* a). *Moreover assume there exists for each* ρ *a measurable map* $\Lambda_\rho: B\to B$ *with finite-dimensional range, such that*

$$\begin{gathered} E\Lambda_\rho X_1=0, \quad E\|X_1-\Lambda_\rho X_1\|^{2+\delta}\leqq\rho^{2+\delta}, \\ E\|n^{-1/2}\sum_{j\geqq n}(X_j-\Lambda_\rho X_j)\|^2\leqq\rho^2. \end{gathered} \tag{4.3}$$

Then the conclusion of part a) *holds.*

Remark. The assumption in a) are exactly the same as in Theorem 1 of Dehling and Philipp (1982), where an almost sure invariance principle is proved.

Proof. b) We first show that the sequence Z_n converges in distribution to some probability measure on $C_B[0,1]$. If π denotes the Prohorov distance on the space of all measures on $C_B[0,1]$, it is sufficient to show that the distributions of Z_n form a Cauchy sequence with respect to π. Let Z_n^ρ be the random element of $C_B[0,1]$ defined as in (4.2) with X_j replaced by $\Lambda_\rho X_j$. Then we know by Theorem 4 of Kuelbs and Philipp (1980) that for fixed ρ the sequence (Z_n^ρ) converges in distribution as $n \to \infty$. If we can show that $\pi(\mathscr{L}(Z_n^\rho), \mathscr{L}(Z_n))$ converges to zero uniformly in n as $\rho \to 0$, we are done.

From (4.3) and Lemma 3.3 we know that

$$E\left\| n^{-1/2} \sum_{j=1}^{n} (X_j - \Lambda_\rho X_j) \right\|^{2+\alpha} \leqq C\rho^{2+\alpha}$$

for some constant C, where $\alpha = \varepsilon\delta/8$. Hence by stationarity

$$E\left\| \sum_{j=k}^{n} (X_j - \Lambda_\rho X_j) \right\|^{2+\alpha} \leqq C\rho^{2+\alpha} |n-k|^{1+\alpha/2}.$$

If we define $M_n^\rho = \max_{k \leqq n} \| \sum_{j \leqq k} (X_j - \Lambda_\rho X_j) \|$, then Theorem 12.2 in Billingsley (1968), which continues to hold for Banach space valued random variables, implies that for all $\lambda > 0$

$$P\{M_n^\rho > \lambda\} \leqq K \lambda^{-(2+\alpha)} n^{1+\alpha/2} \rho^{2+\alpha}.$$

If we put $\lambda = K^{\frac{1}{3+\alpha}} \cdot n^{1/2} \rho^{\frac{2+\alpha}{3+\alpha}}$ we get

$$P\{M_n^\rho > n^{1/2} K^{\frac{1}{3+\alpha}} \rho^\gamma\} \leqq K^{\frac{1}{3+\alpha}} \rho^\gamma \qquad \textit{where } \gamma = (2+\alpha)(3+\alpha)^{-1}.$$

Since $\|Z_n - Z_n^\rho\| \leqq n^{-1/2} M_n^\rho$, this implies that $P(\|Z_n - Z_n^\rho\| > C\rho^\gamma) \leqq C\rho^\gamma$ for some constant C and hence

$$\pi(\mathscr{L}(Z_n), \mathscr{L}(Z_n^\rho)) \leqq C\rho^\gamma.$$

To finish the proof, we have to show that the only possible limit point for the sequence $\{Y_n\}$ is Brownian motion with covariance structure T. But this can be done by a minor modification of the proof of the corresponding part of Theorem 3.1 of Eberlein (1979). Note that in his proof of this part strong mixing is sufficient and that all the other assumptions are easily verified, since we work with a sequence of random variables rather than with an array. Part a) follows from part b) in the usual way.

5. Proof of Theorem 1, Part b)

First we prove a special case of Theorem 1 where B is the d-dimensional Euclidean space and where X_i are independent. We state this result separately.

Proposition 5.1. *Let $X_1, \ldots, X_n$ be independent $\mathbb{R}^d$-valued random variables with $EX_i = 0$, $E\|X_i\|^{2+\delta} < \infty$. If μ_n and ν_n are defined as in Theorem B, then*

$$\pi(\mu_n, \nu_n) \leqq c\, n^{-\delta/8} d^{1/4} \rho_{2+\delta}^{1/4} (1 + |\log(n^{-\delta/2} d^{-1} \rho_{2+\delta})|^{1/2}) \tag{5.1}$$

where $\rho_{2+\delta} = \frac{1}{n} \sum_{i=1}^{n} E\|X_i\|^{2+\delta}$ and c is again an absolute constant.

Proof. Without loss of generality we can assume that $n^{-\frac{1}{2}\delta}\rho_{2+\delta}<1$. Define: $Y_i=X_i 1_{\{\|X_i\|\leqq\sqrt{n}\}}$, $\beta_i=EY_i$, $Z_i=Y_i-\beta_i$. Then we have:

$$E\|Z_i\|^3\leqq 4E\|Y_i\|^3=4E\|X_i\|^3 1_{\{\|X_i\|\leqq\sqrt{n}\}} \leqq 4E\|X_i\|^{2+\delta} n^{\frac{1}{2}(1-\delta)}$$

$$\rho_3=\frac{1}{n}\sum_{i=1}^{n} E\|Z_i\|^3\leqq 4n^{\frac{1}{2}(1-\delta)}\rho_{2+\delta}.$$

Thus we have by (1.8):

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{i=1}^{n} Z_i\right);\, N(0,\operatorname{cov}(n^{-1/2}\sum_{i=1}^{n} Z_i))\right)$$
$$\leqq C d^{1/4}\rho_{2+\delta}^{1/4} n^{\frac{1-\delta}{8}} n^{-1/8}(1+|\log(d^{-1} n^{\frac{1-\delta}{2}} n^{-1/2}\rho_{2+\delta})|^{1/2})$$
$$=C d^{1/4}\rho_{2+\delta}^{1/4} n^{-\delta/8}(1+|\log(d^{-1} n^{-\delta/2}\rho_{2+\delta})|^{1/2}). \tag{5.2}$$

Since

$$E\|X_i-Z_i\|^2\leqq E\|X_i-Y_i\|^2=E\|X_i\|^2 1_{\{\|X_i\|\geqq\sqrt{n}\}} \leqq n^{-\delta/2} E\|X_i\|^{2+\delta}$$

we have:

$$E\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i-\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i\right\|^2\leqq\left(n^{-1}\sum_{i=1}^{n} E\|X_i\|^{2+\delta}\right) n^{-\delta/2} =\rho_{2+\delta} n^{-\delta/2}.$$

Using this, Čebyšev's inequality and Lemma 2.1 we get:

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{i=1}^{n} X_i\right);\,\mathscr{L}\left(n^{-1/2}\sum_{i=1}^{n} Z_i\right)\right)\leqq n^{-\delta/6}\rho_{2+\delta}^{1/3}, \tag{5.3}$$

$$\pi\left(N\left(0,\operatorname{cov}\left(n^{-1/2}\sum_{i=1}^{n} X_i\right)\right);\, N\left(0,\operatorname{cov}\left(n^{-1/2}\sum_{i=1}^{n} Z_i\right)\right)\right)\leqq n^{-\delta/6}\rho_{2+\delta}^{1/3}. \tag{5.4}$$

Now using (5.2), (5.3), (5.4) we get the desired result.

Remark. Instead of (5.1) we will often use the following inequality.

$$\pi(\mu_n,\nu_n)\leqq C n^{-\delta/9} d^{1/3}\rho_{2+\delta}^{2/9}. \tag{5.5}$$

Now we conclude the proof of Theorem 1.

First we shall prove (1.16) and (1.18). Since all d-dimensional inner-product space are isometrically isomorphic, we may as well assume that $B=\mathbb{R}^d$. Then we define for every $n\in\mathbb{N}$:

$$p(n)=[n^{1/2}],\qquad q(n)=[n^{1/3}],\qquad l(n)=\left[\frac{n}{p+q}\right] \tag{5.6}$$

and we get

$$n^{1/2}\ll l(n)\ll n^{1/2}. \tag{5.7}$$

For the moment let n be fixed. We define blocks H_j, $1\leqq j\leqq l$, and I_j, $1\leqq j\leqq l+1$ of consecutive integers such that:

$$\begin{aligned}&\operatorname{card} H_j=p, \quad \operatorname{card} I_j=q \qquad 1\leqq j\leqq l\\&\operatorname{card} I_{l+1}=n-l(p+q)\end{aligned} \tag{5.8}$$

and the order of the blocks is $H_1, I_1, H_2, \ldots, H_l, I_l, I_{l+1}$. We further define

$$\begin{aligned}X_j&=\sum_{\nu\in H_j}\xi_\nu \qquad 1\leqq j\leqq l\\Z_j&=\sum_{\nu\in I_j}\xi_\nu \qquad 1\leqq j\leqq l+1.\end{aligned} \tag{5.9}$$

Hence $\sum_{j=1}^{n}\xi_j=\sum_{j=1}^{l}X_j+\sum_{j=1}^{l+1}Z_j$ and therefore using Lemma 3.5 we get

$$\begin{aligned}&E\left\|n^{-1/2}\sum_{j=1}^{l}X_j-n^{-1/2}\sum_{j=1}^{n}\xi_j\right\|^2\\&\quad=n^{-1}E\left\|\sum_{j=1}^{l+1}Z_j\right\|^2\ll n^{-1}(lq+p+q)\rho_{2+\delta}^{2/(2+\delta)}\ll n^{-1/6}\rho_{2+\delta}^{2/(2+\delta)}\end{aligned}$$

which implies

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^{l}X_j\right), \mathscr{L}\left(n^{-1/2}\sum_{j=1}^{n}\xi_j\right)\right)\ll n^{-1/18}(\rho_{2+\delta}^{1/3}+1). \tag{5.10}$$

By Theorem 2 of Berkes and Philipp [2] we can redefine the sequence $(X_j, 1\leqq j\leqq l)$ on a new probability space together with a sequence $(Y_j, 1\leqq j\leqq l)$ of independent random variables such that X_j and Y_j for each j have the same distribution and such that:

$$P\{\|X_j-Y_j\|\geqq 6\phi(q)\}\leqq 6\phi(q).$$

Hence

$$\pi\left(L\left(n^{-1/2}\sum_{j=1}^{l}X_j\right), L\left(n^{-1/2}\sum_{j=1}^{l}Y_j\right)\right)\leqq 6l\,\phi(q)\ll n^{-1/6}. \tag{5.11}$$

Now we are going to apply Proposition 5.1 to the sequence $\{l^{1/2}n^{-1/2}Y_j, 1\leqq l\leqq j\}$. By Lemma 3.5 and 3.6 and by definition of l and Y_j we know that

$$E\|l^{1/2}n^{-1/2}Y_j\|^{2+\delta}\ll\rho_{2+\delta}. \tag{5.12}$$

If we denote by T_n' the covariance of $l^{1/2}n^{-1/2}Y_j$, then we get using (5.5)

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{j=1}^{l}Y_j\right), N(0,T_n')\right)\ll l^{-\delta/9}d^{1/3}\rho_{2+\delta}^{2/9}=n^{-\delta/18}d^{1/3}(\rho_{2+\delta}^{1/3}+1). \tag{5.13}$$

We finally have to estimate $\pi(N(0, T_n'); N(0, T))$. This will be done in two steps. First we get using Lemma 2.1:

$$\begin{aligned}\pi(N(0, T_n'); N(0, T_p)) &\leqq (E\|Y_1(p^{-1/2} - l^{1/2}n^{-1/2})\|^2)^{1/3}\\ &\ll (n^{-1/2}(n^{1/2} - (pl)^{1/2}))^{2/3}(E\|p^{-1/2}Y_1\|^2)^{1/3}\\ &\ll (n^{-1}(lq+p+q))^{2/3}\rho_{2+\delta}^{2/3(2+\delta)^{-1}}\\ &\ll n^{-1/9}(1+\rho_{2+\delta}^{1/3}). \end{aligned} \tag{5.14}$$

Using (3.13) and Theorem 7 we finally get

$$\pi(N(0, T_p), N(0, T)) \ll d^{1/4} p^{-\varepsilon/4} \sigma^{1/2} \ll d^{1/4} n^{-\varepsilon/8}(1+\rho_{2+\delta}^{1/3}). \tag{5.15}$$

Note that in all the inequalities above the constants implied by $\ll$ depend only on ε, δ and the constant implied by $\ll$ in (1.14). Now put (5.10), (5.11), (5.13)–(5.15) together and get (1.16) and (1.18).

The general result is now an easy consequence of (1.16) and the fact that the Banach-Mazur distance of a d-dimensional Banach space B from $\mathbb{R}^d$ does not exceed d.

Lemma 5.1. *Let B be a d-dimensional Banach space with norm $\|\cdot\|$. Then there exists an inner product norm $\|\cdot\|_1$ on B such that*

$$\frac{1}{d}\|x\|_1 \leqq \|x\| \leqq \|x\|_1 \qquad x \in B. \tag{5.16}$$

Proof. See Lindenstrauss, Tzafriri (1977), p. 17, proof of Lemma 1.c.4.

Now let $\|\cdot\|_1$ be the inner product norm defined above on B and let π_1, π be the Prohorov metrics on $\mathscr{M}(B)$ with respect to the norms $\|\cdot\|_1$, $\|\cdot\|$. Then it is easy to see that

$$\frac{1}{d}\pi_1(\cdot,\cdot) \leqq \pi(\cdot,\cdot) \leqq \pi_1(\cdot,\cdot).$$

Hence we have from (1.16)

$$\pi\left(\mathscr{L}\left(n^{-1/2}\sum_{i=1}^{n}\xi_i\right); N(0, T)\right) \leqq C n^{-\lambda_1} d^{1/3}(1+(\rho_{2+\delta}^{(1)})^{1/3})$$

where $\rho_{2+\delta}^{(1)} = \sup_{i \geqq 1} E\|\xi_i\|_1^{2+\delta} \leqq d^{2+\delta}\rho_{2+\delta}$ and this implies (1.15).

6. Proof of Theorem 1, Part a)

Because of the lack of an approximation theorem like Theorem A for the strong mixing case, we have to use a different method to prove part a) of Theorem 1. Our proof is essentially a variant of Yurinskii's proof of Theorem B.

Proposition 6.1. *Let $X_1, \ldots, X_n$ be a set of $\mathbb{R}^d$-valued random variables with $EX_i = 0$ and $\rho_3 := \sup E\|X_i\|^3 < \infty$. Let $\mathfrak{M}_a^b$ denote the σ-field generated by $X_a, \ldots, X_b$ and define $\alpha = \sup \alpha(\mathfrak{M}_1^k, \mathfrak{M}_{k+1}^n)$. Let $Z_1, \ldots, Z_n$ be a set of independent $N(0, \operatorname{cov} X_i)$-distributed random variables. If $f\colon \mathbb{R}^d \to \mathbb{R}$ is three times differentiable, then*

$$|Ef(n^{-1/2}(X_1+\ldots+X_n)) - Ef(n^{-1/2}(Z_1+\ldots+Z_n))| \leq C(\alpha^{2/3} n^{1/2} \sup_x \|D_x f\| \rho_3^{1/3} + d^2 \alpha^{1/3} \sup_x \|D_x^2 f\| \rho_3^{2/3} + n^{-1/2} \sup_x \|D_x^3 f\| \rho_3 \tag{6.1}$$

where C is an absolute constant.

Proof. We can take $Z_1, \ldots, Z_n$ to be independent of $X_1, \ldots, X_n$. Then

$$\begin{aligned}
&|Ef(n^{-1/2}(X_1+\ldots+X_n)) - Ef(n^{-1/2}(Z_1+\ldots+Z_n))| \\
&\quad \leq |Ef(n^{-1/2}(X_1+\ldots+X_n)) - Ef(n^{-1/2}(X_1+\ldots+X_{n-1}+Z_n)) \\
&\qquad + \ldots \\
&\qquad + |Ef(n^{-1/2}(X_1+\ldots+X_i+Z_{i+1}+\ldots+Z_n)) \\
&\qquad\quad - Ef(n^{-1/2}(X_1+\ldots+X_{i-1}+Z_i+\ldots+Z_n)) \\
&\qquad + \ldots \\
&\qquad + |Ef(n^{-1/2}(X_1+Z_2+\ldots+Z_n)) - Ef(n^{-1/2}(Z_1+\ldots+Z_n))|.
\end{aligned}$$

We estimate each of the n summands by using Taylor-expansion of f about $n^{-1/2}(X_1+\ldots+X_{i-1}+Z_{i+1}+\ldots+Z_n)$. For the sake of brevity denote $W_i = n^{-1/2}(X_1+\ldots+X_{i-1}+Z_{i+1}+\ldots+Z_n)$.

$$\begin{aligned}
&|Ef(n^{-1/2}(X_1+\ldots+X_{i-1}+X_i+Z_{i+1}+\ldots+Z_n)) \\
&\quad - Ef(n^{-1/2}(X_1+\ldots+X_{i-1}+Z_i+Z_{i+1}+\ldots+Z_n))| \\
&\leq |E(D_{W_i} f(n^{-1/2} X_i))| + |ED_{W_i} f(n^{-1/2} Z_i)| \\
&\quad + |ED_{W_i}^2 f(n^{-1/2} X_i) - ED_{W_i}^2 f(n^{-1/2} Z_i)| \\
&\quad + \tfrac{1}{6} \sup_x \|D_x^3 f\| \, E\left(\left\|\frac{1}{\sqrt{n}} X_i\right\|^3 + \left\|\frac{1}{\sqrt{n}} Z_i\right\|^3\right).
\end{aligned}$$

We estimate now these summands separately:

$$D_{W_i} f(n^{-1/2} X_i) = n^{-1/2}(\operatorname{grad} f(W_i); X_i).$$

If $\mathfrak{Z}$ denotes the σ-field generated by $Z_1, \ldots, Z_n$, then $\mathfrak{Z}$ is independent of $\mathfrak{M}_1^n$ and hence

$$\alpha(\mathfrak{M}_1^{i-1} \vee \mathfrak{Z}; \mathfrak{M}_i^i) \leq \alpha.$$

Now W_i is $\mathfrak{M}_1^{i-1} \vee \mathfrak{Z}$-measurable. Hence by Lemma 3.1, we have,

$$|ED_{W_i} f(n^{-1/2} X_i)| \leq 15 \alpha^{2/3} n^{-1/2} \sup_x \|D_x f\| \, \|X_1\|_3$$

and $ED_{W_i} f(n^{-1/2} Z_i) = 0$, since Z_i is independent of W_i.

$$|ED^2_{W_i} f(n^{-1/2} X_i) - ED^2_{W_i} f(n^{-1/2} Z_i)| = \frac{1}{n} |EX_i A(W_i) X_i^T - EZ_i A(W_i) Z_i^T|$$

where $A = A(W_i)$ is a symmetric $d \times d$-matrix $A(W_i) = (a_{k,l})_{1 \leqq k,l \leqq d}$ and where $a_{k,l}$ are realvalued random variables, measurable with respect to $\sigma(W_i)$.

If $X_i = (X_i^{(1)}, \ldots, X_i^{(d)})$, then

$$\begin{aligned}|EX_i A(W_i) X_i^T - EZ_i A(W_i) Z_i^T| &\leqq \sum_{k,l} |E(X_i^{(k)} X_i^{(l)} a_{k,l}) - E(Z_i^{(k)} Z_i^{(l)} a_{k,l})| \\ &\leqq \sum_{k,l} |E(X_i^{(k)} X_i^{(l)} a_{kl}) - E(X_i^{(k)} X_i^{(l)}) E a_{kl}| \\ &\quad + \sum_{k,l} |E(Z_i^{(k)} Z_i^{(l)} a_{kl}) - E(Z_i^{(k)} Z_i^{(l)}) E a_{kl}|\end{aligned}$$

because X_i and Z_i have the same covariance. The second sum vanishes again, because Z_i is independent of W_i. By applying Lemma 3.1. to each term in the first sum, we get:

$$|EX_i A(W_i) X_i^T - EZ_i A(W_i) Z_i^T| \leqq d^2\, 15 \alpha^{1/3} \|X_i\|_3 |\sup_{k,l} \|a_{k,l}\|_\infty$$

and hence

$$|ED^2_{W_i} f(n^{-1/2} X_i) - ED^2_{W_i} f(n^{-1/2} Z_i)| \leqq 15 n^{-1} d^2 \alpha^{1/3} \|X_i\|_3 \sup_x \|D_x^2 f\|.$$

Using the fact that there exists a universal constant c such that $E\|Z_i\|^3 \leqq c(E\|Z_i\|^2)^{3/2}$, we get finally

$$E(\|n^{-1/2} X_i\|^3 + \|n^{-1/2} Z_i\|^3) \leqq c n^{-3/2} \|X_i\|_3^3.$$

Combining the above estimates we obtain the desired result. □

By applying Proposition 6.1. to Yurinskii's smooth "indicators" of sets we can estimate the Prohorov-distance of $\mathscr{L}(n^{-1/2}(X_1 + \ldots + X_n))$ and $N\left(0, \frac{1}{n} \sum_{i=1}^{n} \operatorname{cov} X_i\right)$. Using Lemma 2.3 and Lemma 2.4 we get for every $\varepsilon \in (0, 1)$

$$\begin{aligned}&\pi\left(\mathscr{L}(n^{-1/2}(X_1 + \ldots + X_n)), N\left(0, n^{-1} \sum_{i=1}^{n} \operatorname{cov} X_i\right)\right) \\ &\quad \leqq c(\alpha^{2/3} n^{1/2} \varepsilon^{-1} d^{-1/2} \rho_3^{1/3} + d^2 \alpha^{1/3} \rho_3^{2/3} \varepsilon^{-2} d^{-1/2} + n^{-1/2} \rho_3 \varepsilon^{-3} d^{-1/2}) \\ &\quad + 4\varepsilon + 4\varepsilon \left(\log \frac{1}{\varepsilon}\right)^{1/2} d^{1/2} + 4\varepsilon d^{1/2}.\end{aligned} \tag{6.2}$$

Proposition 6.2. *Let $\{\xi_\nu, \nu \geqq 1\}$ be a strong mixing sequence of $\mathbb{R}^d$-valued random variables centered at expectations and with $(3+\delta)$-th moments uniformly bounded by $\rho_{3+\delta}$ for some $\delta > 0$. Suppose that the mixing coefficients $\alpha(k)$ satisfy* (3.16).

If T_n denotes the covariance operator of $n^{-1/2} \sum_{\nu \leqq n} \xi_\nu$, then

$$\pi(\mathscr{L}(n^{-1/2} \sum_{\nu \leqq n} \xi_\nu); N(0, T_n)) \ll n^{-1/20} d^{3/2}(1+\rho_{3+\delta}^{3/(3+\delta)}) \tag{6.3}$$

where the constant implied by $\ll$ only depends on ε, δ and the constant implied by $\ll$ in (3.16).

Proof. For fixed n put $p=[n^{1/2}], q=[n^{1/4}]$, $l=\left[\frac{n}{p+q}\right]$; $n^{1/2} \ll l \ll n^{1/2}$. Again we define blocks H_j, $1 \leqq j \leqq l$ and I_j, $1 \leqq j \leqq l+1$ of consecutive integers with:

$$\operatorname{card} H_j = p \ (1 \leqq j \leqq l), \operatorname{card} I_j = q \ (1 \leqq j \leqq l), \operatorname{card} I_{l+1} = n - l(p+q).$$

Define $X_j = p^{-1/2} \sum_{\nu \in H_j} \xi_\nu$, $1 \leqq j \leqq l$. Using Lemma 3.2 and Lemma 3.9 we get

$$E\|X_j\|^3 \leqq C b^3, \quad \text{where } b = \rho_{3+\delta}^{1/(3+\delta)}.$$

C is a constant, depending only on ε, δ and the constant implied by $\ll$ in (3.16).

Using Lemma 3.2., Minkowski's inequality and the mean-value theorem we obtain,

$$\begin{aligned}
\left\| l^{-1/2} \sum_{j=1}^{l} X_j - n^{-1/2} \sum_{\nu=1}^{n} \xi_\nu \right\|_2 &\leqq \left\| (l p)^{-1/2} \sum_{j=1}^{l} \sum_{\nu \in H_j} \xi_\nu - n^{-1/2} \sum_{\nu} \xi_\nu \right\|_2 \\
&\leqq n^{-1/2} \left\| \sum_{j=1}^{l+1} \sum_{\nu \in I_j} \xi_\nu \right\|_2 + ((l p)^{-1/2} - n^{-1/2}) \left\| \sum_{j=1}^{l} \sum_{\nu \in H_j} \xi_\nu \right\|_2 \\
&\ll (n^{-1/2}(l q + p + q)^{1/2} + (l p)^{1/2}(n^{-1/2} - (l p)^{-1/2}))\, b \\
&\ll (n^{-1/2}(n^{3/4} + n^{1/2})^{1/2} + n^{-1}(n - l p))\, b \\
&\ll (n^{-1/8} + n^{-1}(n^{3/4} + n^{1/2}))\, b \\
&\ll n^{-1/8}\, b.
\end{aligned}$$

Hence:

$$\pi\left(\mathscr{L}\left(l^{-1/2} \sum_{j=1}^{l} X_j\right); \mathscr{L}\left(n^{-1/2} \sum_{\nu=1}^{n} \xi_\nu\right)\right) \ll n^{-1/12} b^{2/3}$$

and by Lemma 2.1

$$\pi(N(0, \operatorname{cov}(l^{-1/2} \sum_{j \leqq l} X_j)); N(0, T_n)) \ll n^{-1/12} b^{2/3}.$$

Next we use (6.2) to estimate the Prohorov-distance of $\mathscr{L}(l^{-1/2} \sum_{j \leqq l} X_j)$ and $N(0, l^{-1} \sum_{j \leqq l} \operatorname{Cov} X_j)$. Note that $\alpha(k) \ll k^{-3}$, hence $\alpha = \alpha(q) \ll n^{-3/4}$. If we put $\varepsilon = n^{-1/15}$, we get:

$$\pi(\mathscr{L}(l^{-1/2}\sum_{j\leq l}X_j);N(0,l^{-1}\sum_{j\leq l}\operatorname{Cov}X_j))$$
$$\ll n^{-1/2}n^{1/4}n^{1/15}d^{-1/2}b+n^{-1/4}n^{2/15}d^{3/2}b^2+n^{-1/4}n^{3/15}d^{-1/2}b^3+n^{-1/15}$$
$$+n^{-1/15}(\log n^{1/15})^{1/2}d^{1/2}+n^{-1/15}d^{1/2}$$
$$\ll n^{-1/20}d^{3/2}(1+b^3).$$

For the sake of brevity we define:

$$T=\operatorname{Cov}(l^{-1/2}\sum_{j\leq l}X_j)\qquad S=l^{-1}\sum_{j\leq l}\operatorname{Cov}X_j.$$

We want to use Theorem 7 to estimate the distance of $N(0,T)$ and $N(0,S)$, so first we have to estimate $\|T-S\|_1$. Let $e_1,\dots,e_d$ be any orthonormal basis for $\mathbb{R}^d$.

$$(Te_k,e_k)=E(l^{-1/2}\sum_{j\leq l}X_j,e_k)^2=l^{-1}\sum_{j\leq l}E(X_j,e_k)^2+l^{-1}\sum_{i\neq j}E(X_i,e_k)(X_i,e_k)$$
$$(Se_k,e_k)=l^{-1}\sum_{j\leq l}E(X_j,e_k)^2.$$

Then

$$|(T-S)e_k,e_k)|\leq 2l^{-1}\sum_{1\leq i<j\leq l}|E(X_i,e_k)(X_j,e_k)|$$
$$\ll l^{-1}\sum_{1\leq i<j\leq l}\bar{\alpha}^{1/3}(j-i)b^2$$
$$=l^{-1}\sum_{i=1}^{l}\bar{\alpha}^{1/3}(i)(l-i)b^2$$

where $\bar{\alpha}$ denotes the mixing coefficient for the set $X_1,\dots,X_l$. Thus $\bar{\alpha}(1)\leq\alpha(q)\ll n^{-3/4}$ and $\bar{\alpha}(i)\leq\alpha((i-1)p)\ll i^{-3}n^{-3/2}$ for $i>1$. Hence

$$|((T-S)e_k,e_k)|\ll l^{-1}n^{-1/4}l\,b^2+l^{-1}\sum_{i=2}^{l}i^{-1}n^{-1/2}(l-i)b^2\ll n^{-1/4}b^2$$

and therefore $\|T-S\|_1\ll dn^{-1/4}b^2$.

Using Theorem 7 we get

$$\pi(N(0,\operatorname{Cov}(l^{-1/2}\sum_{j\leq l}X_j));\quad \mathrm{N}(0,l^{-1}\sum_{j\leq l}\operatorname{Cov}X_j))$$
$$\ll d^{1/4}n^{-1/16}b^{1/2}d^{1/4}=d^{1/2}n^{-1/16}b^{1/2}.\quad\square$$

Now we conclude the proof of Theorem 1, part a.

First we prove (1.13) and (1.17). Then (1.12) can again be deduced from these relations using Lemma 5.1. We may also again assume that $B=R^d$. For fixed n define a sequence $\tau_\nu=\tau_\nu^{(n)}$ by

$$\bar{\tau}_\nu=\xi_\nu 1_{\{\|\xi_\nu\|\leq n^{1/50}\}}\quad\text{and}\quad \tau_\nu=\bar{\tau}_\nu-E\bar{\tau}_\nu$$

$\{\tau_\nu, \nu \geqq 1\}$ is strong mixing with the same mixing coefficients as $\{\xi_\nu, \nu \geqq 1\}$ and has finite 12-th moments. Since $\delta \leqq 2/3$, we have:

$$\alpha(k) \ll k^{-(1+\varepsilon)\cdot 4} = k^{-(1+\varepsilon)(3+9/9)}.$$

Hence $\{\tau_\nu, \nu \geqq 1\}$ satisfies the assumptions made in Proposition 6.2.

$$E\|\tau_i\|^{12} \leqq 2^{11} E\|\xi_i\|^{12} 1_{\{\|\xi_i\| \leqq n^{1/50}\}} \leqq E\|\xi_i\|^{2+\delta} n^{(10-\delta)/50}$$

$$(E\|\tau_i\|^{12})^{3/12} \leqq (E\|\xi_i\|^{2+\delta})^{1/4} n^{(10-\delta)1/200} \leqq \rho_{2+\delta}^{1/4} n^{(10-\delta)1/200}.$$

If S_n denotes the covariance operator of $n^{-1/2} \sum_{\nu \leqq n} \tau_\nu$, then we get by (6.3)

$$\pi(\mathscr{L}(n^{-1/2} \sum_{\nu \leqq n} \tau_\nu); N(0, S_n)) \ll n^{-1/20} d^{3/2}(1 + \rho_{2+\delta}^{1/4} n^{1/20-\delta/200})$$
$$\leqq n^{-\delta/200} d^{3/2}(1 + \rho_{2+\delta}^{1/4}).$$

Now we have to estimate the error introduced by replacing $\{\xi_\nu, \nu \geqq 1\}$ by $\{\tau_\nu, \nu \geqq 1\}$.

If $\delta' = \delta \dfrac{2}{2+\varepsilon}$, then $1 + 2/\delta' < (1+\varepsilon)(1+2/\delta)$, so that there exist an $\varepsilon' > 0$, such that

$$(1+\varepsilon')(1+2/\delta') < (1+\varepsilon)(1+2/\delta)$$

and hence $\alpha(k) \ll k^{-(1+\varepsilon')(1+2/\delta')}$.

Then $E\|\xi_i - \tau_i\|^{2+\delta'} \leqq 4E\|\xi_i\|^{2+\delta} n^{-(\delta-\delta')/50} \leqq 4\rho_{2+\delta} n^{-\delta\bar{\varepsilon}/50}$ where $\bar{\varepsilon} = \varepsilon(2+\varepsilon)^{-1}$.
Hence by (3.3)

$$E\|n^{-1/2} \sum_{\nu \leqq n} (\xi_\nu - \tau_\nu)\|^2 \leqq C \rho_{2+\delta}^{2/(2+\delta')} n^{-1/50\delta\bar{\varepsilon}2/(2+\delta')}.$$

Using Čebyšev's inequality and Lemma 2.1 this implies,

$$\pi(\mathscr{L}(n^{-1/2} \sum_{\nu \leqq n} \xi_\nu); \mathscr{L}(n^{-1/2} \sum_{\nu \leqq n} \tau_\nu)) \leqq C \rho_{2+\delta}^{1/3} n^{-1/200\delta\bar{\varepsilon}}$$

and $\pi(N(0, T_n); N(0, S_n)) \leqq C \rho_{2+\delta}^{1/3} n^{-1/200\delta\bar{\varepsilon}}$.

Using 3.14 and Theorem 7 we finally get:

$$\pi(N(0, T_n), N(0, T)) \leqq C d^{1/2} n^{-\varepsilon/4} \rho_{2+\delta}^{1/(4+2\delta)}.$$

Putting everything together gives (1.13) and (1.17).

7. Proof of Theorem 2

We shall prove Theorem 2 for both cases simultaneously.

We introduce the following notation:

$$\mu_n = \mathscr{L}\left(n^{-1/2} \sum_{j=1}^{n} X_j\right), \mu_n^N := \mu_n \circ P_N^{-1} = \mathscr{L}\left(n^{-1/2} \sum_{j=1}^{n} P_N \circ X_j\right)$$

$\mu^N = w - \lim_{n \to \infty} \mu_n^N$, where w-lim means the limit in the weak topology on $\mathscr{M}(B)$.

This limit exists by Theorem 1. By (1.19) and Čebyšev's inequality

$$\pi(\mu_n^N, \mu_n) \ll N^{-s/3}.$$

By Theorem 1 we have with $\rho_{2+\delta}^N = \sup_j E\,\|P_N X_j\|^{2+\delta}$

$$\begin{aligned}\pi(\mu_n^N, \mu^N) &\ll n^{-\lambda} N^3 \ (1+(\rho_{2+\delta}^N)^{1/3}) \\ &\ll n^{-\lambda} N^{3+r} \ (1+\rho_{2+\delta}^{1/3}).\end{aligned}$$

Now let $m \geqq n$ and take $N = [n^\gamma]$ where $\gamma = \dfrac{3\lambda}{9+3r+s}$.

Hence

$$\begin{aligned}\pi(\mu_n, \mu_m) &\leqq \pi(\mu_n, \mu_n^N) + \pi(\mu_n^N, \mu^N) + \pi(\mu^N, \mu_m^N) + \pi(\mu_m^N, \mu_m) \\ &\ll n^{-\frac{s\gamma}{3}} + n^{-\lambda} n^{\gamma(3+r)} \ (1+\rho_{2+\delta}^{1/3}) \\ &\ll n^{-\kappa} \ (1+\rho_{2+\delta}^{1/3}).\end{aligned}$$

Hence $(\mu_n, n \geqq 1)$ is a Cauchy sequence in $(\mathscr{M}(B), \pi)$ and since this is a complete metric space, $\mu = \lim \mu_n$ exists. That μ is a Gaussian measure with covariance operator T defined by (1.11) is easily seen by application of Theorem 1 to the one-dimensional sequence $(f \circ X_j, j \geqq 1)$, $f \in B^*$.

Hence we get

$$\pi(\mu_n, N(0, T)) = \lim_m \pi(\mu_n, \mu_m) \ll n^{-\kappa}(1+\rho_{2+\delta}^{1/3})$$

which proves (1.21).

8. Proofs of Theorem 3 and Corollary 1

We shall arrange the proof of Theorem 3 in such a way that we prove both cases (ϕ-mixing and absolutely regular) simultaneously. To prove Theorem 3 all we need is that the sequence $\{X_j, j \geqq 1\}$ is absolutely regular with $\beta(k) \ll k^{-2(1+\varepsilon)}$, that $EX_j = 0$, $E\|X_j\|^{2+\delta} \leqq \rho_{2+\delta}$, that (1.21) is satisfied and that for some constant $C < \infty$

$$E\left\|\sum_{j=a+1}^{a+n} X_j\right\|^{2+\alpha} \leqq C n^{1+\alpha/2} \quad \text{for all } a \geqq 0, \quad n \geqq 1 \tag{8.0}$$

where $\alpha = \dfrac{\varepsilon\delta}{8}$.

To prove (8.0) we use in the absolutely regular case Lemma 3.3 together with (1.19) and Lemma 3.2 and in the ϕ-mixing case Lemma 3.6 together with (1.19) and Lemma 3.5.

Let κ be as in (1.21) and define

$$\lambda = 2 + \kappa^{-1}, \qquad n_k = [k^{\lambda-1}], \qquad \gamma = \frac{\varepsilon\delta\kappa}{256(1+2\kappa)}. \tag{8.1}$$

Let H_k and I_k be blocks of consecutive integers such that

$$\operatorname{card} H_k = n_k \qquad \operatorname{card} I_k = [k^2]. \tag{8.2}$$

The order of the blocks is $H_1, I_1, H_2, I_2, \ldots$. Furthermore define

$$t_{k+1} = \sum_{j=1}^{k} \operatorname{card}(H_j \cup I_j). \tag{8.3}$$

From this definition we get using (8.1) and (8.2)

$$k^\lambda \ll t_k \ll k^\lambda. \tag{8.4}$$

If we define $Y_k = n_k^{-1/2} \sum_{v \in H_k} X_v$, then we have using (1.21)

$$\pi(\mathscr{L}(Y_k); N(0, T)) \leqq C n_k^{-\kappa}$$

where C is a constant. Now Y_k and Y_{k+1} are separated by the block I_k and hence we can apply Theorem A to get a sequence $\{Z_k, k \geqq 1\}$ of independent $N(0, T)$-distributed B-valued random variables such that:

$$\begin{aligned} P\{\| Y_k - Z_k\| &\geqq 2 C n_k^{-\kappa} + 2\beta^{1/2}(\operatorname{card} I_{k-1})\} \\ &\leqq 2 C n_k^{-\kappa} + 2\beta^{1/2}(\operatorname{card} I_{k-1}) \\ &\ll k^{-\kappa(\lambda-1)} + k^{-(2+2\varepsilon)} \ll k^{-1-\kappa} + k^{-(2+2\varepsilon)}. \end{aligned} \tag{8.5}$$

By the Borel-Cantelli Lemma and using the fact that $\kappa < 1/8$ we get then:

$$\| Y_k - Z_k\| \ll n_k^{-\kappa}. \tag{8.6}$$

If $X(t)$ is any B-valued Brownian motion with covariance structure T then $(t_{k+1} - t_k)^{-1/2}(X(t_{k+1}) - X(t_k))$ has the same distribution as Y_k. Hence using the same kind of argument as Kuelbs and Philipp (1980), p. 1024 we can assume without loss of generality that there exists a Brownian motion $X(t)$ with covariance structure T such that

$$(t_{k+1} - t_k)^{-1/2}(X(t_{k+1}) - X(t_k)) = Z_k. \tag{8.7}$$

We shall show that $X(t)$ satisfies (1.22). First note that by Fernique's Lemma ([10])

$$\|Z_k\| \ll \log k \quad \text{a.s.} \quad \text{as } k \to \infty.$$

Using this, (8.2), (8.3), (8.6) and (8.7) we get

$$\begin{aligned} \Big\| \sum_{v \in H_k} X_v - (X(t_{k+1}) - X(t_k))\Big\| &= \|n_k^{1/2} Y_k - (t_{k+1} - t_k)^{1/2} Z_k\| \\ &\ll n_k^{1/2} \| Y_k - Z_k\| + |n_k^{1/2} - (t_{k+1} - t_k)^{1/2}| \, \|Z_k\| \\ &\ll n_k^{1/2-\kappa} + n_k^{-1/2} k^2 \log k \\ &\ll n_k^{1/2-\kappa}. \end{aligned}$$

Summing these inequalities we then get

$$\left\|\sum_{j=1}^{k}\sum_{\nu\in H_j} X_\nu - X(t_{k+1})\right\| \ll \sum_{j=1}^{k} n_j^{1/2-\kappa} \ll k^{(1/2-\kappa)(\lambda-1)+1} \ll t_k^{(1/2-\kappa)(1-1/\lambda)+1/\lambda} \ll t_k^{1/2-\gamma} \tag{8.8}$$

since $(1/2-\kappa)(1-1/\lambda)+1/\lambda = 1/2-\kappa/2 < 1/2-\gamma$ by (8.1).

Lemma 8.1. $\sum_{j=1}^{k} \|\sum_{\nu\in I_j} X_\nu\| \ll t_k^{3/\lambda}$.

Proof. By Čebyšev's inequality and (8.0) we have:

$$P\{\|\sum_{\nu\in I_k} X_\nu\| \geqq k^2\} \leqq k^{-2(2+\alpha)} E\|\sum_{\nu\in I_k} X_\nu\|^{2+\alpha} \ll k^{-2(2+\alpha)} k^{2\left(1+\frac{\alpha}{2}\right)} \ll k^{-2}.$$

Hence again by the Borel-Cantelli Lemma we have

$$\|\sum_{\nu\in I_k} X_\nu\| \ll k^2$$

which implies $\sum_{j=1}^{k} \|\sum_{\nu\in I_j} X_\nu\| \ll k^3 \ll t_k^{3/\lambda}$.

Lemma 8.2. $\max_{t_k<t\leqq t_{k+1}} \left\|\sum_{\nu=t_k+1}^{t} X_\nu\right\| \ll t_k^{1/2-\gamma}$.

This can be proved in almost the same way as Proposition 2.2 in Kuelbs and Philipp (1980). The only difference is that instead of their Lemma 2.5 we use (8.0).

Using Fernique's Lemma [10] it is easy to show that

$$\max_{t_k\leqq t\leqq t_{k+1}} \|X(t)-X(t_k)\| \ll t_k^{1/2-\gamma}. \tag{8.9}$$

Now let $t>0$ and choose k such that $t_k<t\leqq t_{k+1}$. Then we get using Lemma 8.1, Lemma 8.2, (8.8) and (8.9)

$$\|\sum_{\nu\leqq t} X_\nu - X(t)\| \leqq \left\|\sum_{\nu=t_k+1}^{t} X_\nu\right\| + \|X(t)-X(t_k)\| + \|\sum_{j<k}\sum_{\nu\in H_j} X_\nu - X(t_k)\| + \sum_{j<k}\|\sum_{\nu\in I_j} X_\nu\| \ll t_k^{1/2-\gamma}. \quad \square$$

We now prove Corollary 1. First consider the sequence $X_\nu^* = X_\nu - L(X_\nu)$, $\nu\geqq 1$ and define $P_N x = \sum_{\nu\leqq N}(X, e_\nu)e_\nu$. The range of P_N has dimension N and (1.20) is satisfied with $r=0$. We shall show that (1.19) holds for the sequence $\{X_\nu^*, \nu\geqq 1\}$. If $\{X_\nu, \nu\geqq 1\}$ is ϕ-mixing, then we have by Lemma 3.5 and (1.25) that

$$E\left\|n^{-1/2}\sum_{\nu=a+1}^{a+n}(X_\nu^* - P_N X_\nu^*)\right\|^2 \leqq E\|X_1^* - P_N X_1^*\|^2((1+\sum_{j\geqq 1}\phi^{1/2}(j)) \ll \sum_{\nu>N} E(X_1, e_\nu)^2 \ll N^{-s}.$$

If $\{X_\nu, \nu \geqq 1\}$ is absolutely regular, this remains valid because of Lemma 3.2 and (1.25′). Hence we conclude that (1.21) and Theorem 3 both hold for $\{X_\nu^*, \nu \geqq 1\}$. Now it is easy to see that the covariance function T_L of the sequence $\{L(X_\nu), \nu \geqq 1\}$ vanishes identically. Next we use the following proposition, whose proof we shall postpone until the end of the section.

Proposition 8.1. *Let $\{Y_j, j \geqq 1\}$ be a strictly stationary sequence of H-valued random variables centered at expectations and with finite $(2+\delta)$-th moments with $0<\delta \leqq 1$. Moreover suppose that $\{Y_j, j \geqq 1\}$ is either ϕ-mixing or absolutely regular with the usual mixing rates* (1.22) *resp.* (1.14) *and that the covariance function T of the sequence $\{Y_j, j \geqq 1\}$ which is defined as in* (1.11) *vanishes for all $f, g \in H$. Then we have*

$$E \|\sum_{j \leqq n} Y_j\|^2 \ll n^{1-\varepsilon}. \tag{8.10}$$

Now we conclude the proof of the corollary. By (8.10) we have

$$E \|n^{-1/2} \sum_{j \leqq n} (X_j - X_j^*)\|^2 \ll n^{-\varepsilon}$$

and hence $\pi(\mathscr{L}(n^{-1/2} \sum_{j \leqq n} X_j); \mathscr{L}(n^{-1/2} \sum_{j \leqq n} X_j^*)) \ll n^{-\varepsilon/3}$. This combined with the facts that (1.21) holds for $\{X_j^*, j \geqq 1\}$ and that $\{X_j, j \geqq 1\}$ and $\{X_j^*, j \geqq 1\}$ both have the same covariance operator proves that (1.21) holds. To prove that the conclusion of Theorem 3 holds note that all we needed in the proof of it were the absolute regularity with $\beta(k) \ll k^{-(2+\varepsilon)}$, the fact that $EX_j = 0$, $E\|X_j\|^{2+\delta} < \infty$, (1.21) and relation (8.0), which by Lemma 3.2 resp. 3.5 always hold. This concludes the proof of Corollary 1.

The proof of Proposition 8.1 follows the lines of the proof of Propisition 4.3 of Kuelbs and Philipp (1980). First note that their Lemma 4.9 continues to hold in our situation, hence

$$\sigma_y^2 = \lim_{n \to \infty} n^{-1} E \|\sum_{j \leqq n} Y_j\|^2 = 0.$$

Using Lemma 3.1 resp. Lemma 3.4 we deduce by standard arguments that $E \|\sum_{j \leqq n} Y_j\|^2 = n\sigma_y^2 + O(n^{1-\varepsilon})$, which proves (8.10).

9. Proof of Theorem 4

Our first Lemma gives us an estimate for the speed of convergence in the central limit theorem of the kind needed for the proof of the law of the iterated logarithm.

Lemma 9.1. *Under the assumptions of Theorem 4, there exists a Gaussian measure μ, such that*

$$\pi(\mathscr{L}(n^{-1/2} \sum_{j \leqq n} X_j); \mu) \ll (\log N)^{-\alpha} \tag{9.1}$$

where α is the constant from (1.24).

The proof for this is almost the same as the proof of Theorem 2 and we shall omit it. The following maximal inequality is due to Berkes (1974).

Lemma 9.2. *Let* $\{X_j, j \geqq 1\}$ *be a strong mixing sequence of B-valued random variables with* $EX_i = 0$, $\sup E\|X_j\|^{2+\delta} < \infty$. *Suppose that the mixing coefficients* $\alpha(k)$ *satisfy* (1.23). *Let* σ^2 *be defined by* $\sigma^2 = \sup_{a,n} E\|n^{-1/2}(X_{a+1} + \ldots + X_{a+n})\|^2$. *Then there exist constants* $\rho > 0$ *and* $n_0 \in \mathbb{N}$ *such that for all* $n \geqq n_0$

$$P\{\max_{k \leqq n} \|S_k\| > x\} \leqq 2P\{\|S_n\| > x - 20n^{1/2}\sigma\} + n^{-\rho}. \tag{9.2}$$

Proof. We follow the lines of a similar proof in Reznik (1968) and Berkes (1974).

Put $A_k = \{\|S_1\| \leqq x, \ldots, \|S_{k-1}\| \leqq x, \|S_k\| > x\}$ to get,

$$P\{\max_{1 \leqq k \leqq n} \|S_k\| > x\} = \sum_{k=1}^{n} P(A_k).$$

Let $\rho > 0$ be a fixed number (to be chosen later) and denote by $\sum' P(A_k)$ the sum over those k for which $P(A_k) > n^{-(1+\rho)}$. Hence to prove (9.2) it suffices to show that $\sum' P(A_k) \leqq 2P\{\|S_n\| > x - 20\sigma n^{1/2}\}$. By a standard argument, this will follow if we show

$$P(\|S_n - S_k\| > 20\sigma n^{1/2} \mid A_k) \leqq \tfrac{1}{2}$$

for those k's for which $P(A_k) > n^{-(1+\rho)}$. For every $p \geqq 1$ we have

$$\begin{aligned} P(\|S_n - S_k\| > 20\sigma n^{1/2} \mid A_k) \leqq & P(\|X_{k+1} + \ldots + X_{k+p}\| > 10\sigma n^{1/2} \mid A_k) \\ & + P(\|X_{k+p+1} + \ldots + X_n\| > 10\sigma n^{1/2} \mid A_k). \end{aligned}$$

Using Minkowski's inequality the first probability is bounded from above by

$$n^{1+\rho} n^{-(1+\delta/2)} \sigma^{-(2+\delta)} 10^{-(2+\delta)} p^{2+\delta} \sup_{j \geqq 1} E\|X_j\|^{2+\delta} \leqq C n^{\rho - \delta/2} p^{2+\delta}.$$

The second probability is bounded by

$$\begin{aligned} & P(\|X_{k+p+1} + \ldots + X_n\| \geqq 10\sigma n^{1/2}) + \frac{\alpha(p)}{P(A_k)} \\ & \leqq 10^{-2}\sigma^{-2} n^{-1} E\|X_{k+p+1} + \ldots + X_n\|^2 + p^{-(2+\varepsilon)(1+2/\delta)} n^{1+\rho} \\ & \leqq 10^{-2} + p^{-(2+\varepsilon)(1+2/\delta)} n^{1+\rho}. \end{aligned}$$

Take a $\beta > 0$ such that

$$\delta(2+\delta)^{-1}(2+\varepsilon)^{-1} < \beta < \tfrac{1}{2}\delta(2+\delta)^{-1}.$$

Then we can find a $\rho > 0$ such that

$$(2+\delta)\beta < \delta/2 - \rho \quad \text{and} \quad 1 + \rho < \beta(2+\varepsilon)(1+2/\delta).$$

Take $p=[n^\beta]$. Then

$$P(\|S_n-S_k\|>20\,\sigma\,n^{1/2}|A_k)\leqq C\,n^{\beta(2+\delta)+\rho-\delta/2}+10^{-2}+n^{1+\rho-\beta(2+\varepsilon)(1+2/\delta)}$$
$$\leqq \tfrac{1}{2}$$
for n big enough. □

Our next step is to prove a bounded law of the iterated logarithm. By Fernique's theorem there exists a constant c which depends only on the space B such that for all Gaussian measures ν we have

$$\int \exp\frac{\|x\|^2}{c^2\,\sigma^2}\,d\nu(x)\leqq 1$$

where $\sigma^2=\int \|x\|^2\,d\nu(x)$.

Proposition 9.1. *Let $\{X_j, j\geqq 1\}$ be a weakly stationary strong mixing sequence of B-valued random variables, centered at expectations and with finite $(2+\delta)$-th moments for some $0<\delta<\frac{2}{3}$. Suppose that the mixing coefficients satisfy* (1.23). *Moreover suppose that there exists a Gaussian measure μ and a constant $\alpha>1$ such that*

$$\pi(\mathscr{L}(n^{-1/2}\sum_{j\leqq n} X_j);\mu)\ll(\log n)^{-\alpha}.$$

Then, if σ^2 is defined as in Lemma 9.2 *and $a_n=(2\,n\log\log n)^{1/2}$*

$$\limsup_{n\to\infty} a_n^{-1}\sum_{j\leqq n} X_j\leqq 4\,c\,\sigma.$$

Proof. By a well-known theorem we have: $\int \|x\|^2\,d\mu\leqq\sigma^2$. Define $n_k=e^k$, $S_n=\sum_{\nu\leqq n} X_\nu$

$$P\left\{\overline{\lim}\,\frac{\|S_n\|}{a_n}>4\,c\,\sigma\right\}\leqq P\left\{\sup_{n_k<n\leqq n_{k+1}}\frac{\|S_n\|}{a_n}>4\,c\,\sigma\ \text{i.o. in}\ k\right\}$$
$$\leqq P\{\sup_{n\leqq n_{k+1}}\|S_n\|>4\,c\,\sigma\,a_{n_k}\ \text{i.o. in}\ k\}.$$

For k so large that $n_k\geqq n_0$, where n_0 is as in Lemma 9.2, we obtain

$$\begin{aligned}
&P\{\sup_{n\leqq n_{k+1}}\|S_n\|>4\,c\,\sigma\,a_{n_k}\}\\
&\leqq 2P\{\|S_{n_{k+1}}\|>4\,c\,\sigma\,a_{n_k}-20\,\sigma\sqrt{n_{k+1}}\}+n_{k+1}^{-\rho}\\
&\ll P\{\|n_{k+1}^{-1/2}\,S_{n_{k+1}}\|>2\,c\,\sigma\sqrt{2\log k}\}+k^{-2}\\
&\ll \mu\{x|\,\|x\|>2\,c\,\sigma\sqrt{2\log k}-c(\log k)^{-\alpha}\}+k^{-\alpha}\\
&\ll \mu\{x|\,\|x\|>2\,c\,\sigma(\log k)^{1/2}\}+k^{-\alpha}\\
&\ll \exp-((2\,c\,\sigma(\log k)^{1/2})^2\,c^{-2}\,\sigma^{-2})+k^{-\alpha}\\
&\ll k^{-\alpha}.
\end{aligned}$$

Since $\alpha>1$, the Borel-Cantelli Lemma shows that $\overline{\lim}\,a_n^{-1}\,\|S_n\|\leqq 4\,c\,\sigma$.

Proof of Theorem 4. Following Theorem 2 of [6] all we have to show is a compact law of the iterated logarithm. By the law of the iterated logarithm for finite-dimensional spaces we know that $\{a_n^{-1} \sum_{v \leqq n} P_N X_v, n \geqq 1\}$ is relatively compact almost surely for each N. Since $(I - P_d)$ is a bounded linear map, the estimation for the π-distance (9.1) holds also for the sequence $\{X_v - P_N X_v, v \geqq 1\}$. If σ_N is defined by

$$\sigma_N^2 = \sup_n E \|n^{-1/2} \sum_{v \leqq n} (X_v - P_N X_v)\|^2$$

then by the previous proposition

$$\overline{\lim}\, a_n^{-1} \sum_{v \leqq n} (X_v - P_N X_v) \leqq 4 \sigma_N C.$$

Since $\sigma_N \to 0$ as $N \to \infty$, we conclude by standard arguments that $a_n^{-1} \sum_{v=1}^{n} X_v$ is a.s. relatively compact.

10. Proof of Theorem 5

We define a sequence $(e_n)_{n \geqq 1}$ in $C[0,1]$ by

$$e_n(t) = \begin{cases} \cos 2\pi k t & \text{if } n = 2k-1 \\ \sin 2\pi k t & \text{if } n = 2k. \end{cases}$$

Note that $e_n' = (2\pi n)^{-1} e_n \in \Lambda_\alpha$. As usual, an inner product on $C[0,1]$ is given by $(f, g) = \frac{1}{2} \int f(t) g(t)\, dt$.

Lemma 10.1. *There exists a sequence $(\lambda_n) = (\lambda_n^\alpha)$ and a constant $C = C_\alpha$ such that*

$$\sum_{n \geqq 1} \lambda_n^{-2} < \infty, \tag{10.1}$$

$$\sum_{n \geqq N} \lambda_n^2 ((f, e_{2n})^2 + (f, e_{2n-1})^2) \leqq C N^{\frac{1}{2} - \alpha} \quad \textit{for all } f \in \Lambda_\alpha, \tag{10.2}$$

$$\sum_{n \geqq 1} |(f, e_n)| \leqq C \quad \textit{for all } f \in \Lambda_\alpha. \tag{10.3}$$

Proof. The proof of (10.3) can be found in Zygmund (1935), p. 135. For the proof of (10.1) and (10.2) let $\beta = 1/2(\alpha + 1/2)$ and $\lambda_n = n^\beta$. Since $\alpha > 1/2$ we have that (10.1) holds. From Zygmund (1935), p. 136 (3) we get that

$$\sum_{2^{v-1} < n \leqq 2^v} ((f, e_{2n-1})^2 + (f, e_{2n})^2) \leqq \bar{C}\, 2^{-2v\alpha}$$

for some constant $\bar{C}$ depending only on α. Hence

$$\sum_{2^{v-1} < n \leqq 2^v} \lambda_n^2 ((f, e_{2n-1})^2 + (f, e_{2n})^2) \leqq \bar{C}\, 2^{2\beta v} 2^{-2\alpha v} = \bar{C}\, 2^{(\frac{1}{2} - \alpha) v}.$$

If $N \in \mathbb{N}$ choose k such that $2^{k-1} < N \leqq 2^k$ and then

$$\sum_{n \geqq N} \lambda_n^2((f, e_{2n})^2 + (f, e_{2n-1})^2) \leqq \sum_{n > 2^{k-1}} \lambda_n^2((f, e_{2n})^2 + (f, e_{2n-1})^2)$$
$$\leqq \bar{C} \sum_{v \geqq k} 2^{v(\frac{1}{2}-\alpha)} \leqq C N^{\frac{1}{2}-\alpha}. \quad \square$$

Recall that in order to prove Theorem 5 all we have to do is to find the mappings P_N with the properties (1.19) and (1.20).

For $x \in C(\Lambda_\alpha)$ define $P_{2N}(x)$ by

$$P_{2N} x(f) = \sum_{n \leqq 2N} 2\pi n\, x(e_n')(f, e_n)$$

so that we have:

$$P_{2N} X_j(f) = \sum_{n \leqq 2N} e_n(\xi_j)(f, e_n).$$

It is easily seen that it is not necessary to define P_{2N-1}.

i) $P_{2N}(C(\Lambda_\alpha))$ is spanned by the $2N$ mappings $f \to (f, e_n), n = 1, \ldots, 2N$, hence $\dim P_{2N}(C(\Lambda_\alpha)) \leqq 2N$.

ii) First note that Theorem 2 remains valid if (1.20) is replaced by the following weaker statement:

$$\|P_{2N} X_j\| \leqq C N^r \|X_j\| \quad \text{a.s. for some constant } C. \tag{10.4}$$

Now (10.4) can be proved as follows:

$$\|P_{2N} X_j\| = \sup_{f \in \Lambda_\alpha} |\sum_{n \leqq 2N} e_n(\xi_j)(f, e_n)| \leqq \sup_{f \in \Lambda_\alpha} \sum_n |(f, e_n)| \ll 1.$$

iii)
$$\left\| \sum_{j=1}^n (X_j - P_{2N} X_j) \right\|^2 = \sup_{f \in \Lambda_\alpha} \left| \sum_{j=1}^n (X_j(f) - P_{2N} X_j(f)) \right|^2$$
$$\leqq \sup_{f \in \Lambda_\alpha} \left| \sum_{j=1}^n \left(f(\xi_j) - \sum_{k=1}^{2N} (f, e_k)\, e_k(\xi_j) \right) \right|^2$$
$$\leqq \sup_{f \in \Lambda_\alpha} \left| \sum_{j=1}^n \sum_{k > 2N} (f, e_k)\, e_k(\xi_j) \right|^2$$
$$\leqq \sup_{f \in \Lambda_\alpha} \left| \sum_{k > 2N} \sum_{j=1}^n (f, e_k)\, e_k(\xi_j) \right|^2$$
$$\leqq 2 \sup_{f \in \Lambda_\alpha} \left| \sum_{k > N} \sum_{j=1}^n (f, e_{2k-1})\, e_{2k-1}(\xi_j) \right|^2$$
$$+ 2 \sup_{f \in \Lambda_\alpha} \left| \sum_{k > N} \sum_{j=1}^n (f, e_{2k})\, e_{2k}(\xi_j) \right|^2.$$

Now we estimate the first and the second summand separately.

The first summand is bounded by:

$$\leqq \sup_{f\in A_\alpha}\left|\sum_{k>N}(f,e_{2k-1})\lambda_k\sum_{j=1}^{n}\lambda_k^{-1}e_{2k-1}(\xi_j)\right|^2$$

$$\leqq \sup_{f\in A_\alpha}\left(\sum_{k>N}(f,e_{2k-1})^2\lambda_k^2\right)\left(\sum_{k>N}\lambda_k^{-2}\left(\sum_{j=1}^{n}e_{2k-1}(\xi_j)\right)^2\right)$$

$$\ll N^{\frac{1}{2}-\alpha}\sum_{k>N}\lambda_k^{-2}\left(\sum_{j=1}^{n}e_{2k-1}(\xi_j)\right)^2.$$

A similar estimate holds for the second summand, hence

$$E\left\|\sum_{j=1}^{n}(X_j-P_NX_j)\right\|^2\ll N^{\frac{1}{2}-\alpha}\sum_{k>N}\lambda_k^{-2}\left(E\left(\sum_{j=1}^{n}e_{2k-1}(\xi_j)\right)^2+E\left(\sum_{j=1}^{n}e_{2k}(\xi_j)\right)^2\right)$$

$$\ll N^{\frac{1}{2}-\alpha}\sum_{k>N}\lambda_k^{-2}Cn$$

$$\ll N^{\frac{1}{2}-\alpha}\cdot n.$$

The second step follows from calculations similar to those in Lemma 3.2 and Lemma 3.5.

11. Proof of Theorem 6.

First we need some facts from topology.

Definition. a) Let X be a topological space. A set $\{\varphi_i, i\in I\}$ of real-valued functions on S is called a partition of unity if

i) The supports of the φ_i form a neighborhood-finite closed covering of X.

ii) $0\leqq\varphi_i(x)\leqq 1$ for all $x\in X$, all $i\in I$.

iii) $\sum_{i\in I}\varphi_i(x)=1$.

b) If $\{U_i, i\in I\}$ is an open covering of Y, we say that a partition $\{\varphi_i, i\in I\}$ of unity is subordinated to $\{U_i, i\in I\}$ if the support of φ_i lies in the corresponding U_i.

The following lemma is a special case of Theorem 4.2 in Dugundji (1966).

Lemma 11.1. *If S is a compact space and $\{U_i, i\in I\}$ a finite open covering of S, then there exists a partition of unity subordinated to $\{U_i, i\in I\}$.*

Of course to prove Theorem 6 we want to apply Theorem 2 and 3. Hence again we have to define operators P_N: $C(S)\to C(S)$. Let $N\geqq 1$ be given. Then we can cover S with N balls of radius $g(N)$ by assumption. Let $\{s_i,\ 1\leqq i\leqq N\}$ be the centers of these balls and let $\{\varphi_i,\ 1\leqq i\leqq n\}$ be a partition of unity subordinated to this covering.

Define for $x\in C(S)$

$$P_Nx=\sum_{i=1}^{N}x(s_i)\varphi_i.$$

Clearly the range of P_N has dimension N and using the fact that $\sum_{i \leqq N} \varphi_i = 1$ we get that $\|P_N\| = 1$. Moreover $(x - P_N x)(s) = \sum_{i \leqq N} (x(s) - x(s_i))\varphi_i$. If $s \in S$, then $\varphi_i(s) = 0$ unless $d(s, s_i) \leqq g(N)$. Hence we see that

$$\|x - P_N x\| \leqq \sup\{|x(s) - x(s')| : d(s, s') \leqq g(N)\}.$$

This together with (1.30) and (1.31) shows that P_N satisfies (1.19) which finishes the proof of Theorem 6.

Acknowledgement. This research was done during a stay at the University of Illinois. I am grateful to Walter Philipp for his hospitality and his constant advice during this time.

References

1. Berkes, I.: Almost sure invariance principles for mixing processes. Unpublished manuscript (1974)
2. Berkes, I., Philipp, W.: Approximation theorems for independent and weakly dependent random vectors. Ann. Probability. **7**, 29–54 (1979)
3. Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968
4. Cartan, H.: Differential Calculus. Paris: Hermann 1971
5. Davydov, Yu.A.: The invariance principle for stationary processes. Theory Probability Appl. **14**, 487–498 (1970)
6. Dehling, H., Philipp, W.: Almost sure invariance principles for weakly dependent vector-valued random variables. Ann. Probability **10**, 689–701 (1982)
7. Doob, J.L.: Stochastic Processes. New York: Wiley 1953
8. Dugundji, F.: Topology. Boston: Allyn and Bacon 1966
9. Eberlein, E.: An invariance principle for lattices of dependent random variables. Z. Wahrscheinlichkeitstheorie verw. Gebiete **50**, 119–133 (1979)
10. Fernique, X.: Intégrabilité des vecteurs Gaussiens. C.R. Acad. Sci. Paris **270**, 1698–1699 (1970)
11. Ibragimov, I.A.: Some limit theorems for stationary processes. Theory Probability Appl. **7**, 349–382 (1962)
12. Ibragimov, I.A., Rosanov, Yu.A.: Gaussian Random Processes. Berlin-Heidelberg-New York: Springer 1978
13. Kaufman, R., Philipp, W.: A uniform law of the iterated logarithm for classes of functions. Ann. Probability **6**, 930–952 (1978)
14. Kuelbs, J., Philipp, W.: Almost sure invariance principles for partial sums of mixing B-valued random variables. Ann. Probability. **8**, 1003–1036 (1980)
15. Kuo, H.H.: Gaussian Measures in Banach Spaces. Lecture Notes in Mathematics **463**. Berlin-Heidelberg-New York: Springer 1975
16. Lindenstrauss, J., Tzafriri, L.: Classical Banach Spaces I. Berlin-Heidelberg-New York: 1977
17. Marcus, M.B., Philipp, W.: Almost sure invariance principles for sums of B-valued random variables with applications to random Fourier series and the empirical characteristic process. Trans. Amer. Math. Soc. **269**, 67–90 (1982)
18. Philipp, W.: Almost sure invariance principles for sums of B-valued random variables. In: Probability in Banach Spaces II. Lecture Notes in Mathematics **709**, 171–193. Berlin-Heidelberg-New York: Springer 1979
19. Philipp, W., Stout, W.F.: Almost sure invariance principles for partial sums of weakly dependent random variables. Mem. Amer. Math. Soc. **161** (1975)
20. Reznik, M.Kh.: The law of the iterated logarithm for some classes of stationary processes. Theory Probability Appl. **13**, 606–621 (1968)

21. Sotres, D.A., Ghosh, M.: Strong convergence of linear rank statistics for mixing processes. Sankhya Ser. B **39**, 1–11 (1977)
22. Trotter, H.F.: An elementary proof of the central limit theorem. Arch. Math. **10**, 226–234 (1959)
23. Volkonskii, V.A., Rozanov, Yu.A.: Some limit theorems for random functions I. Theory Probability Appl. **4**, 178–197 (1959)
24. Yurinskii, V.V.: On the error of the Gaussian approximation for convolutions. Theory Probability Appl. **22**, 236–247 (1977)
25. Zygmund, A.: Trigonometrical series. Warsawa-Lwow 1935

Received May 17, 1981; in revised form January 7, 1983