The Annals of Probability
1986, Vol. 14, No. 3, 922–942

ON THE RATE OF CONVERGENCE IN THE CENTRAL LIMIT THEOREM IN BANACH SPACES

BY F. GÖTZE

University of Bielefeld

Let E denote a separable Banach space and let X_i, $i \in \mathbb{N}$, be a sequence of i.i.d. E-valued random vectors having finite third moment such that the central limit theorem holds. We prove that the convergence rate in the central limit theorem is $O(n^{-1/2})$ for regions $\{x \in E: F(x) < r\}$ which are defined by means of a smooth real valued function F on E, provided that the limiting distribution of the gradient of F fulfills a variance condition.

Using this result we prove that the rate of convergence in the functional limit theorem for empirical processes is of order $O(n^{-1/2})$.

1. Introduction and results. Let E denote a separable Banach space normed by $\| \; \|$. Assume that X_i, $i \in \mathbb{N}$, is a sequence of i.i.d. E-valued random vectors. Let Q denote the common probability distribution of the X_i's defined on the Borel σ-field, say $\mathscr{B}$, of E. Let Q_n denote the probability distribution of $R_n = n^{-1/2}(X_1 + \cdots + X_n)$. Suppose that for some $\gamma \geq 0$

$$E\|X_1\|^{3+\gamma} < \infty \quad \text{and} \quad EX_1 = 0. \tag{1.1}$$

Furthermore, assume that the central limit theorem holds for Q_n, i.e.,

(1.2) Q_n, $n \in \mathbb{N}$, converges weakly to some Gaussian Borel measure N defined on $\mathscr{B}$.

The results of Hoffmann-Jørgensen and Pisier (1976) show that the central limit theorem holds for every i.i.d. sequence X_i, $i \in \mathbb{N}$, satisfying $E\|X_1\|^2 < \infty$ and $EX_1 = 0$ if E is a Banach space of type 2 (which includes for example the L^p function space with $2 \leq p < \infty$).

Let F: $E \to \mathbb{R}$ denote a measurable functional. We shall study the speed of convergence in the central limit theorem for regions $\{x \in E: F(x) < r\}$, $r \in \mathbb{R}$, where F is differentiable. From another point of view this means that we study the speed of convergence in the functional limit theorem for $F(R_n)$ with respect to the Kolmogoroff distance of distribution functions.

For this purpose it will be convenient to use Frechét differentiability of F with respect to $\| \cdot \|$, i.e.,

$$|F(x+h) - F(x) - DF(x)[h]| = o(\|h\|),$$

where $h \to DF(x)[h]$ denotes a continuous linear functional on E. The kth derivative of F at x, say $D^kF(x)$, is a k-linear symmetric functional $(h_1, \ldots, h_k) \to D^kF(x)[h_1, \ldots, h_k]$ with finite norm

$$\|D^kF(x)\| = \sup\{|D^kF(x)[h_1, \ldots, h_k]|: \|h_1\|, \ldots, \|h_k\| \leq 1\}.$$

Received June 1981; revised July 1985.

AMS 1980 subject classifications. Primary 60B12; secondary 60F17.

Key words and phrases. Central limit theorem in Banach spaces, functional limit theorems, empirical processes.

Assume the derivatives of F satisfy the

DIFFERENTIABILITY CONDITION D_β.

$$\|D^kF(x)\| \le c_F(1 + \|x\|^p), \qquad 0 \le k \le 3,$$
(1.3)
$$\|D^3F(x) - D^3F(y)\| \le c_F(1 + \|x\|^p + \|y\|^p)\|x - y\|^\beta$$

for every $x, y \in E$ and some fixed constants $c_F > 0$, $p \ge 0$, and $1 \ge \beta > 0$. In this case F is said to be of *class* $C^{3+\beta}$.

Let $v(x) = E(DF(x)[X_1]^2)$. The crucial condition on F and X_1 is the following "variance" condition which guarantees the existence of a smooth density of the distribution of $F(R)$, where R has Gaussian distribution N on E.

VARIANCE CONDITION.
Let $k_\beta = 6 + 4/\beta$. Let $\varphi_{\boldsymbol\eta} = \prod_1^r \|X_j\|^3 I(\eta_j\|X_j\| \le 1)$ and $X_{\boldsymbol\eta} = \eta_1 X_1 + \cdots + \eta_r X_r$, $r = [k_\beta/2] + 2$. (Here $[x]$ denotes the largest integer smaller or equal to x.) Assume there exists $\eta > 0$ such that

$$\text{(1.4)} \quad E\varphi_{\boldsymbol\eta} I\big(v(R + X_{\boldsymbol\eta}) \le \varepsilon\big) = o(\varepsilon^{k_\beta}), \qquad \varepsilon \to 0 \text{ and } |\eta_j| \le \eta,\ 1 \le j \le r.$$

REMARK. When

$$E\|X_1\|^{3+\gamma} < \infty, \qquad \gamma > 0,$$

we may choose $\eta = 0$ in (1.4) and replace (1.4) by

$$\text{(1.4')} \qquad P(v(R) \le \varepsilon) = O(\varepsilon^{k_\beta}), \quad \text{where } k_\beta = 4/\gamma + 4/\beta + 6.$$

These assumptions combine conditions on the covariance structure of X_1 (or R) and conditions on the "regular" behaviour of the region where the gradient of F vanishes, which will be discussed later.

The main result of this paper is

1.5 THEOREM. *Assume that conditions* (1.1)–(1.4) *hold for some* c_F, p, *and* β. *Then*

$$\sup_{r \in \mathbb{R}} |P(F(R_n) \le r) - P(F(R) \le r)| = O(n^{-1/2}).$$

The validity of the variance condition (1.5) is essential for the convergence speed. A class of Hilbert space examples constructed by Rhee and Talagrand (1984) shows that for functions F which are "flat" at some points $[\|DF(x)\| = O(\|x - x_0\|^q)$ for *every* $q > 0]$ and with a sufficiently "singular" covariance structure of X_1 the rate of convergence can be made slower than any prescribed order.

The speed of convergence for functionals $F(x) = \|x\|$ which are differentiable has been investigated first by Kandelaki (1965) [obtaining $O(1/\log n)$], Kuelbs and Kurtz (1974) who obtained $O(n^{-1/8})$ in Hilbert space, and Paulauskas (1976) who improved the rate for three times differentiable norms in Banach space to $O(n^{-1/6})$. See also the survey articles of Paulauskas (1979) and Rachkauskas (1980) and the monograph of Sazonov (1981).

For functionals of "polynomial type" (i.e., the qth derivative of F vanishes identically) the distribution of $F(R_n)$ can be approximated in regular cases by smooth expansions up to the order $O(n^{-q/2+\varepsilon})$ (Götze, 1984, 1985). Similar results for general differentiable functionals F such that the qth derivative satisfies a variance condition like (1.4) will be treated in a forthcoming paper.

In the following we discuss some applications of Theorem 1.5.

1.6 COROLLARY. *Suppose E is a separable Hilbert space and C denotes a measurable set which is star-shaped with respect to $0 \in C$. Assume that*

(i) $E\|X_1\|^{3+\gamma} < \infty$, $\gamma > 0$ *and* $EX_1 = 0$.

(ii) $F(x) = \inf\{r > 0\colon xr^{-1} \in C\}^4$ *satisfies differentiability condition* (1.3).

(iii) $P(R \in \varepsilon C) = O(\varepsilon^{gk_\beta})$, $g > 10$, k_β *as in* (1.4').

(iv) *The (ordered) eigenvalues of the covariance operator of N, say λ_j, are decreasing such that*

$$\sum_{i=m}^{\infty} \lambda_i = o(m^{-\lambda}/\log m) \quad \text{as } m \to \infty,$$

where $\lambda = 10/(g-10)$. *Then*

$$\sup_{r>0} |P(R_n \in rC) - P(R \in rC)| = O(n^{-1/2}). \tag{1.7}$$

1.8 REMARKS. Condition (iii) holds if C is bounded in norm and more than gk_β eigenvalues are different from zero.

Our conjecture is that condition (ii) can be replaced by a Lipschitz condition on F when we restrict the supremum in (1.7) to $r > \delta > 0$.

For norms in l_2 Nagaev and Chebotarev (1978) obtained $O(n^{-1/2})$ for the case of r.v.'s with independent components. The author proved (1979) that for a large class of U-statistics including symmetric quadratic functionals in Hilbert space the actual rate of convergence is $O(n^{-1+\varepsilon})$ in nondegenerate situations. Furthermore, asymptotic expansions are possible. The moment conditions for the norm in Hilbert space have been weakened in papers by Yurinskii (1982) using a similar approach together with strong exponential estimates for sums (Yurinskii, 1976) to the assumption of a third moment for the rate $O(n^{-1/2})$. For the symmetric case Zalesskii (1982) obtained $O(n^{-(1+\delta)/2})$ assuming a moment of order $3+\delta$, $\delta < 1$. Further papers by Sazonov and Zalesskii (1985), Nagaev (1985), and Opsipov and Rotar (1984) focus on weaker variance conditions and the nonidentical case for the norm in Hilbert space.

For a typical application of the functional version of Theorem 1.5 consider the empirical process

$$x_n(t) = n^{-1/2} \sum_{i=1}^{n} (I(X_i \le t) - t)$$

based on an i.i.d. sample $X_1, \dots, X_n$ taken from the uniform distribution on $[0,1]$. Consider functionals of the following additive type,

$$F(x(\cdot)) = \int_0^1 V(t, x(t))\, dt,$$

such that for $0 \le i \le 4$

$$\sup_{0 \le t \le 1} |D_x^i V(t,x)| \le c(1 + |x|^p). \tag{1.9}$$

Furthermore, define $f(x(\cdot), s) = \int_0^1 (I(s \le t) - t) D_x V(t, x(t))\, dt$ and suppose that for a Brownian bridge $w(t)$, $0 \le t \le 1$,

$$P\left(\int_0^1 f(w(\cdot), s)^2\, ds \le \varepsilon\right) = O(\varepsilon^{k_\beta}), \quad k_\beta \text{ as in (1.4)}. \tag{1.10}$$

We have

1.11 COROLLARY. *Suppose that conditions* (1.9) *and* (1.10) *hold. Then*

$$\sup_z |P(F(x_n(\cdot)) < r) - P(F(w(\cdot)) < r)| = O(n^{-1/2}).$$

1.12 REMARK. Condition (1.10) holds, in particular, if

$$|V(t,x) - V(s,x)| \le c|t-s|^{1/2}(1 + |x|^p) \quad \text{for some } p > 0, \tag{1.13}$$

$$P\left(\int_0^1 D_x V(t, w(t))^2\, dt \le \varepsilon\right) = O(\varepsilon^{5.1 k_\beta}), \quad k_\beta \text{ as in (1.4)}. \tag{1.14}$$

1.15 EXAMPLES. It can be shown by means of Remark 1.12 that the order of convergence is $O(n^{-1/2})$ in Corollary 1.11 for the following functions $V(t,x)$:

(i) $V(t,x) = c(t)x^p$, $p \in \mathbb{N}$, $c(t)$ Lipschitz continuous of exponent 1,
(ii) $D_x V(0,0) \ne 0$ or $D_x V(1,0) \ne 0$, $V(t,x)$ satisfies (1.13),
(iii) $D_x V(t,x) \ge 0$ for all $x \in \mathbb{R}$, $t \in [0,1]$ and $P(\int_\delta^{1-\delta} D_x V(t, w(t))^2\, dt \le \varepsilon) = O(\varepsilon^{k_\beta})$ for small $\delta > 0$.

In example (iii) condition (1.10) follows immediately (interchanging integrations) since the covariance kernel of the Brownian bridge is bounded from below by $c(\delta) > 0$ on $(s,t) \in [\delta, 1-\delta] \times [\delta, 1-\delta]$.

The previous result shows that convergence rates $O(n^{-1/2}\log n)$ based on the strong approximation techniques of Komlós, Major, and Tusnády (1975, 1976) can be improved for some classes of functionals.

We believe that for the particular case of the empirical process the conditions (1.9) and (1.10) are still too restrictive.

The method of proof is a kind of "partial integration" scheme for sums of i.i.d. random vectors in E (see Lemma 3.7) which reduces to Stein's method of partial integration (or differential equations) in the case of linear functionals F. In order to guarantee the existence of the terms obtained by partial integration an obvious condition is $Ev(R)^{-q} < \infty$, hence condition (1.4) for $\eta = 0$ [compare (3.23)].

2. Proof of the results.

PROOF OF COROLLARY 1.6. Since the region is star-shaped the functional $M(x) = F(x)^{1/4}$ is homogeneous, i.e., $M(\lambda x) = \lambda M(x)$ for $\lambda > 0$. Hence

$$M(x) = \langle DM(x), x \rangle, \qquad x \neq 0,$$

where $\langle x, y \rangle$ denotes the scalar product in E.

Denote by C the covariance operator of N with nonnegative eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots$ and $\lambda_0 = 0$. Let $\{e_k,\ k = 1, 2, \dots\}$ denote an orthonormal system of eigenvectors of C corresponding to λ_k, $k \in \mathbb{N}$, which spans the supporting subspace of N. Denote by P_m resp. P_m^c the projection on the space $\langle e_1, \dots, e_m \rangle$ resp. $\langle e_1, \dots, e_m \rangle^c$. Let $C_m = P_m C P_m$. When $\lambda_m > 0$ then

$$\begin{aligned} M(x) &= \left\langle C_m^{1/2} DM(x), C_m^{-1/2} P_m x \right\rangle + \left\langle DM(x), P_m^c x \right\rangle \\ &\leq \left\| C_m^{1/2} DM(x) \right\| \left\| C_m^{-1/2} P_m x \right\| + \| DM(x) \| \, \| P_m^c x \|. \end{aligned} \tag{2.1}$$

Since $DM(\lambda x) = DM(x)$, $x \neq 0$, $\lambda > 0$, we have by condition (1.3)

$$\| DF(x) \| = 4 \| DM(x) \| M(x)^3 \leq c_F (1 + \|x\|^p).$$

By definition

$$\begin{aligned} v(x) &= \langle CDF(x), DF(x) \rangle \\ &= 16 M(x)^6 \langle CDM(x), DM(x) \rangle \\ &\geq 16 M(x)^6 \left\| C_m^{1/2} DM(x) \right\|^2. \end{aligned} \tag{2.2}$$

We have

$$\{ v(x) \leq \varepsilon \} = A_\varepsilon \cup B_\varepsilon \cup D_\varepsilon \cup E_\varepsilon,$$

where

$$\begin{aligned} B_\varepsilon &= \{ M(x) \leq \varepsilon^{1/g} \}, \\ D_\varepsilon &= \{ \|C_m^{-1/2} P_m x\|^2 \geq m \log \varepsilon^{-k} \}, \\ E_\varepsilon &= \{ \|x\| \geq \varepsilon^{-1/(gp)} \}, \\ A_\varepsilon &= \{ v(x) \leq \varepsilon \} \setminus (A_\varepsilon \cup B_\varepsilon \cup C_\varepsilon). \end{aligned}$$

In the following arguments write k for k_β.

By assumption $P(R \in B_\varepsilon) = O(\varepsilon^k)$, and using Lemma 3.12 we have $P(R \in E_\varepsilon) = O(\varepsilon^k)$. Furthermore, $P(R \in D_\varepsilon) = P(\eta_1^2 + \cdots + \eta_m^2 \geq m \log \varepsilon^{-k})$, where $\eta_1, \dots, \eta_m$ denote i.i.d. $N(0,1)$-variates. Chebyshev's inequality yields

$$\begin{aligned} P(R \in D_\varepsilon) &\leq \exp\left[-m \left(\log \varepsilon^{-k} \right) / m \right] (1 - 2/m)^{-m/2} \\ &= O(\varepsilon^k). \end{aligned}$$

Hence it remains to prove $P(R \in A_\varepsilon) = O(\varepsilon^k)$. If $x \in A_\varepsilon$ then by (2.1) and (2.2)

$$\varepsilon^{1/g} \leq M(x),$$

$$M(x) \leq c \varepsilon^{1/2 - 3/g} \left(m \log \varepsilon^{-k} \right)^{1/2} + c \|P_m^c x\| \varepsilon^{-3/g} (1 + \varepsilon^{-p/(gp)}).$$

Choosing $m = [\varepsilon^{-1+10/g}]$, this yields a lower bound

$$\|P_m^c x\| \geq c\varepsilon^{1/g}\varepsilon^{3/g}\varepsilon^{p/(2g)} \quad \text{for } \varepsilon \text{ sufficiently small.}$$

By Chebyshev's inequality we obtain (since $g > 10$) from this lower bound

$$P(R \in A_\varepsilon) \leq P(\|P_m^c R\|^2 \geq c\varepsilon^{10/g})$$
$$\leq \exp(-t\varepsilon^{10/g}) \prod_{k=m+1}^{\infty} (1 - 2t\lambda_k)^{-1/2},$$

provided $t > 0$. Choose $t = \varepsilon^{-10/g}\log(\varepsilon^{-k})$. Hence by condition 1.6(iv) and the choice of m

$$P(R \in A_\varepsilon) = o(\varepsilon^k). \qquad \square$$

Proof of Remark 1.8. Let K denote a positive constant such that $x \in C$ implies $\|x\| \leq K$. Then

$$P(R \in \varepsilon C) \leq P(\|R\| \leq K\varepsilon)$$
$$\leq eE\exp(-\|R\|^2 K^{-1}\varepsilon^{-1})$$
$$\leq e\prod_{i=1}^{\infty}(1 + 2\lambda_i/(K\varepsilon))^{-1/2}$$
$$= O(\varepsilon^{gk_\beta}). \qquad \square$$

Proof of Corollary 1.11. Condition (i) guarantees that the functional $F(x(\cdot))$ is well defined for every $x(t) \in L^q([0,1])$, where $q = p + 4$.

By the dominated convergence theorem and Hölder's inequality F admits four Frechét derivatives with respect to the norm of $L^q([0,1])$. Hence the differentiability condition (1.3) holds.

The invariance principle in $C[0,1]$ implies $\lim P(F(x_n(\cdot)) \leq r) = P(F(w(\cdot)) \leq r)$ for all continuity points r. Hence the CLT for L^q [compare the remarks following condition (1.1)] together with the continuity of the limit distribution function of $F(R)$, $R \in L^q$ implies $P(F(w(\cdot)) \leq r) = P(F(R) \leq r)$ for every r. The same remark applies to condition (1.10). $\square$

Proof of Remark 1.12. Let $\delta > 0$, arbitrarily small, and $\alpha > 0$ be determined later. Let A_ε denote the set of sample paths $w(t)$ of the Brownian bridge satisfying for ε sufficiently small

(i) $$\sup_{|t-s| \leq \varepsilon^\alpha} |w(t) - w(s)| \leq \varepsilon^{-\delta+\alpha/2},$$

(ii) $$\sup_{0 \leq t \leq 1} |w(t)| \leq \log(\varepsilon^{-1}),$$

(iii) $$v(w(\cdot)) = \int_0^1 f(w(\cdot), s)^2\, ds \leq \varepsilon.$$

By well-known properties of the Brownian bridge we have

(2.3) $$P(v(w(\cdot)) \leq \varepsilon) \leq P(A_\varepsilon) + O(\varepsilon^{k_\beta}).$$

Define $g(s) = f(w(\cdot), s)$ for $0 \le s < 1$ and extend it by $g(k + s) = g(s)$, $k \in \mathbb{Z}$ outside $[0,1]$. Let χ denote a r.v. independent of $w(t)$, $t \in [0,1]$ with symmetric distribution around zero such that $\chi^2 \le \varepsilon^{2\alpha}$ a.s. and $E\chi^2 \ge c\varepsilon^{2\alpha}$. Then

$$\varepsilon \ge E\left(\int_0^1 g(s)^2\, ds - \int_0^1 g(s+\chi)^2\, ds \middle| w(\cdot)\right) + \int_0^1 \mathrm{Var}(g(s+\chi)|w(\cdot))\, ds. \tag{2.4}$$

Conditioning on χ in the first integral on the r.h.s. of (2.4) shows that this integral vanishes since $g(t)$ is defined to be periodic. If A_ε occurs and $[s, s+\chi]$ does not contain 0 or 1 we have

$$g(s+\chi) = g(s) + Dg(s)\chi + O(\log^p(\varepsilon^{-1})(\varepsilon^{-\delta+\alpha/2} + \varepsilon^{\alpha/2})|\chi|).$$

Hence

$$\begin{aligned}\varepsilon &\ge \int_{\varepsilon^\alpha}^{1-\varepsilon^\alpha} \mathrm{Var}(g(s+\chi)|w(\cdot))\, ds\\ &\ge E\chi^2 \int_{\varepsilon^\alpha}^{1-\varepsilon^\alpha} Dg(s)^2\, ds + O(\varepsilon^{-\delta+5\alpha/2})\log^p(\varepsilon^{-1}).\end{aligned}$$

Choosing $\alpha = 2/5 + \delta/2$ we conclude that A_ε implies

$$\int_{\varepsilon^\alpha}^{1-\varepsilon^\alpha} D_x V(s, w(s))^2\, ds \le \varepsilon^{-\delta+1/5}.$$

Restoring integration to $[0,1]$ this immediately proves Remark 1.12. □

PROOF OF THEOREM 1.5. The first step is to replace X_j by its truncation Z_j at n^α where $\alpha = 3/(6+2\gamma)$. Let

$$Z_j = X_j \quad \text{when } \|X_j\| \le n^\alpha \text{ and } Z_j = 0 \text{ otherwise.}$$

Define

$$S_n = n^{-1/2}(Z_1 + \cdots + Z_n).$$

Furthermore, replace the indicator function $x \to I(x \le z)$ by

$$g_{nz}(x) = P(x \le z + Un^{-1/2}),$$

where U denotes a random variable which is symmetric around 0, $|U| \le 1$ a.e. and which admits a Lebesgue density that is infinitely often differentiable.

We have

$$\begin{aligned}&\sup_z |P(F(R_n) \le z) - P(F(S_n) \le z)|\\ &\qquad \le P(F(R_n) \ne F(S_n)) \le \sum_{k=1}^n P(\|X_k\| > n^\alpha)\end{aligned} \tag{2.5}$$

and

$$\begin{aligned}&P(F(S_n) \le z) - P(F(R) \le z)\\ &\qquad \le P(F(S_n) \le \bar{z} + Un^{-1/2}) - P(F(R) \le \bar{z} + Un^{-1/2}) + I_0,\end{aligned} \tag{2.6}$$

where

$$I_0 = 2 \sup_z P\big(F(R) \in [z - n^{-1/2}, z + n^{-1/2}]\big), \qquad \bar{z} = z + n^{-1/2}.$$

By Lemma 3.13 we have

$$I_0 = O(n^{-1/2}). \tag{2.7}$$

From (2.5) together with Chebyshev's inequality and (2.7) as well as its analogous lower bound it follows that

$$\begin{aligned} &\sup_z |P(F(R_n) \le z) - P(F(R) \le z)| \\ &\quad \le \sup_z |Eg_{nz}(F(S_n)) - Eg_{nz}(F(R))| + I_0 + O(n^{-1/2})E\|X_1\|^{3+\gamma} \\ &\quad = I_1 + O(n^{-1/2}), \quad \text{say.} \end{aligned} \tag{2.8}$$

We are going to expand the difference I_1 by means of derivatives of g_{nz}. Since we want to replace the rth derivative [which near z is unbounded of order $O(n^{r/2})$] by the bounded function g_{nz} using "partial integration" (Lemma 3.7) we have to insert factors of order $1 + O_p(n^{-1/2})$ in I_1. These factors provide us with functions which make the application of Lemma 3.7 possible.

Define

$$\Delta_n = n^{-1/2}(\theta_1 Z_{n+1} + \cdots + \theta_B Z_{n+B}), \tag{2.9}$$

where B is a number independent of n and $\theta_1, \dots, \theta_B$ are i.i.d., uniformly distributed in [0, 1] and independent of X_j, $j \in \mathbb{N}$. In the following, random variables t, τ, μ are understood to be independent and uniform in [0, 1] and independent of X_j, θ_j. Define the factors mentioned above by

$$e(x) = \prod_1^B \Big(e(Z_{n+j}, x) + n^{-\delta}\sigma(x)^{-2}\Big), \qquad x \in E, \tag{2.10}$$

$$e(z, x) = DF(x)[z]^2\sigma(x)^{-2}, \qquad \sigma(x)^2 = n^{-\delta} + E(DF(x)[Z_1])^2 \tag{2.11}$$

for some $0 < \delta < \min(\beta, 1/4)/4$ and moreover $\delta < \gamma/(6 + 2\gamma)$ when $\gamma > 0$.

Returning to (2.8) again we shall write g instead of g_{nz} and remark that the following inequalities hold uniformly for all $z \in \mathbb{R}$. Then

$$\begin{aligned} I_1 &\le |Eg(F(S_n + \Delta_n))e(S_n) - Eg(F(R))| \\ &\quad + |Eg(F(S_n + \Delta_n))e(S_n) - Eg(F(S_n))| \\ &= I_2 + I_3, \quad \text{say.} \end{aligned} \tag{2.12}$$

Since the estimation of the term I_3 is particularly involved we describe it in detail.

ESTIMATION OF I_3. Write $F_1 = F(S_n + \Delta_n)$ and $\varepsilon_n = F(S_n) - F_1$. Then

$$\begin{aligned} I_3 &= Ee(S_n)|g(F_1) - g(F_1 + \varepsilon_n)| \\ &\le Ee(S_n)I(|F_1 - z| \le |\varepsilon_n|) \end{aligned} \tag{2.13}$$

by definition of g. (Note that $g(z + \delta) = 1$ or 0 for $|\delta| \ge n^{-1/2}$.)

Using $(a+b)^p \le 2^{p-1}(a^p+b^p)$, $a, b \ge 0$, we have by condition (1.3)

$$|\varepsilon_n| \le c\big(1 + \|S_n^{(1)}\|^p + \|S_n^{(2)}\|^p\big)\|\Delta_n\|,$$

where $S^{(1)} = n^{-1/2}\Sigma_1^M Z_j$, $S^{(2)} = S_n - S_n^{(1)}$, and $M = [n/2]$. Let $b = B - 2$. We have

$$\|\Delta_n\| \le n^{-1/2}\Big\{\sum{}' \|Z_{n+j}\| + \sum{}'' \|Z_{n+j}\|\Big\},$$

where the sum Σ' extends over $j = 1,\dots,[B/2]$ and Σ'' c ·· $j = [B/2]+1,\dots,B$ and

$$I(|F_1 - z| \le (s_1+s_2)(z_1+z_2)) \le \sum I(|\,{}_1 - z| \le 4s_i z_j), \tag{2.14}$$

where $s_i, z_j \ge 0$ and the summation is over $i = 1,2$ and $j = 1,2$. Hence we obtain

$$I_3 \le 4EE\big(e(S_n)I(|F_1 - z| \le n^{-1/2}A_n)\big|\mathscr{C}\big) \tag{2.15}$$

with

$$A_n = c\big(1 + \|S_n^{(2)}\|^p\big)\sum{}'' \|Z_{n+j}\|$$

and

$$\mathscr{C} = \sigma\big(Z_j,\, n/2 \le j \le n,\, n + B/2 \le j \le n + B\big).$$

In order to apply the partial integration, Lemma 3.7, note that

$$\begin{aligned} e(x) &= \sum_k{}^{*} n^{-\delta k}\sigma(x)^{-2B}\prod_{(j)}{}' DF(x)[z_{n+j}]^2 \\ &= e_1(x) + r_n, \quad \text{say,} \end{aligned} \tag{2.16}$$

where Σ^* denotes the summation over all $(B-k)$-tuples of integers between 1 and B and Π' denotes the product over such $(B-k)$-tuples of indices (j) and $e_1(x)$ resp. r_n denotes the sum from $k = 1$ to $b - 1$ resp. $k = b$ to B. We have by (2.15)

$$\begin{aligned} I_3 &\le cEE\big(e_1(S_n)I(|F_1 - z| \le n^{-1/2}A_n)\big|\mathscr{C}\big) + Er_n \\ &= I_4 + Er_n, \quad \text{say.} \end{aligned} \tag{2.17}$$

Let Σ^{**} denote summation over all k between b and B and Π' as in (2.16). By (2.20) and (2.11) we have

$$\begin{aligned} Er_n &\le c\sum_k{}^{**} n^{-\delta k}E\sigma(S_n)^{-2k}\prod{}' e(Z_{n+j}, S_n) \\ &\le c\max\big\{n^{-\delta k}E\sigma(S_n)^{-2k}: k = b, b+1,\dots,B\big\} \end{aligned} \tag{2.18}$$

$$= O(n^{-1/2}) \tag{2.19}$$

by Lemma 3.15 with $B = q/2 = [I(\gamma < 0)4/\gamma + 4/\beta] + 6$ such that $\delta b > \frac{1}{2}$.

Let

$$G_n(x) = \int_{-\infty}^{x-z} I(|a| \le n^{-1/2}A_n)\,da.$$

While applying the "partial integration" Lemma 3.7 replace F by $F_S(x) = F(x + S_n^{(2)})$ and let $g(y) = G_n(y)$ and

$$\Delta = \Delta_n - \theta_1 n^{-1/2} Z_{n+1}.$$

Furthermore, let

$$H(x) = n^{-\delta k} \sigma\left(x + S_n^{(2)}\right)^{-2k-2} \prod_{j \neq 1}{}' e\left(Z_{n+j}, x + S_n^{(2)}\right).$$

Let $M = [n/2]$ and identify Z_{M+1} in Lemma 3.7 with Z_{n+1}. Hence by appropriate choice of the splitting (2.17) (using the i.i.d. assumption) we can estimate a typical summand $DF(x + S_n^{(2)})[Z_{n+1}]^2 H(x)$ of $e_1(x)$ conditioned on $\mathscr{D} = \sigma(Z_j, M+1 \le j \le n, \Delta)$ using

$$Ee\left(Z_{n+j}, x\right) \le 1 \quad \text{and} \quad \sigma(x)^{-2} n^{-\delta} \le 1. \tag{2.20}$$

Lemma 3.7 yields

$$\begin{aligned} I_4 \le{} & cEG_n(F(S_{n+1} + \Delta))\left|DF(S_{n+1} + \Delta)[S_{n+1}]H(S_{n+1}, \Delta)\right| \\ & + \|a_n\| EG_n(F(S_n + \Delta))(1 + \|S_n\|^q) H(S_n, \Delta) \\ & + EG_n(F(S_{n+\tau} + \Delta))\left|H_1(S_{n+\tau}, \Delta)\left[Z_{n+1}^2\right]\right| \\ & + O(n^{-1/2})\left|E\left(H_2(S_{n+\mu\tau}, \Delta)\right)\left[Z_{n+1}^3\right]\right|. \end{aligned} \tag{2.21}$$

By definition H_1 and H_2 consist of derivatives of F up to the second order and a first derivative of $e_1(x)$. The following estimate for this derivative will be frequently used: [notation of (2.16)]

$$\left|De_1(x)\right| \le c\left(1 + \|x\|^{4p}\right) \sum_k{}^* \|Z_{n+k}\|^2 \sigma(x)^{-2k-4} n^{-\delta k} \prod_{j \neq k}{}' e\left(Z_{n+j}, x\right). \tag{2.22}$$

We have $G_n(x) \le cn^{-1/2} A_n$ for every x (with A_n constant given $\mathscr{D}$) and $\|a_n\| = O(n^{-1/2})$. By expansion of Σ'' in A_n and the i.i.d. assumption, together with bounds for H_1 and H_2, relations (2.21) and (2.20) yield for some $q > 0$:

$$\begin{aligned} I_4 \le{} & \left[O(n^{-1/2}) + o\left(\|a_n\| n^{-1/2}\right)\right] E\sigma(S_n)^{-4}\left(1 + \|S_n\|^q\right)\left(1 + \|Z_{n+B}\|^3\right) \\ & + O(n^{-1/2}) E\left(1 + \|S_n\|^q\right)\left(1 + \|Z_{n+1}\|^2 + \|Z_{n+1}\|^3\right)\sigma\left(S_{n+\tau}\right)^{-4} \\ & + O(n^{-1/2}) E\left(1 + \|S_n\|^q\right)\left(1 + \|Z_{n+1}\|^3 + \|Z_{n+B}\|^2\right)\sigma\left(S_{n+\tau\mu}\right)^{-4} \\ ={} & O(n^{-1/2}) \end{aligned} \tag{2.23}$$

by Remark 3.24 with $q = 4$ using similar arguments as in (2.18). Hence (2.17)–(2.23) yield

$$I_3 = O(n^{-1/2}). \tag{2.24}$$

Recall that

$$I_2 = \left|Eg(F(S_n + \Delta_n))e(S_n) - Eg(F(R))e(R)\right|.$$

Let $W_{nt} = S_n \sin(\pi t/2) + R\cos(\pi t/2)$. We may rewrite the difference in I_2 as an integral over the derivative of the function $t \to g(F(\Delta_n \sin(\pi t/2) + W_{nt}))e(W_{nt})$

and obtain with Proposition 3.1(i) and the notation used there,

$$
\begin{aligned}
I_2 &\le |EDg(F(S_{nt}))DF(S_{nt})[T_{nt}]e(W_{nt})| \\
&\quad + |Eg(F(S_{nt}))De(W_{nt})[T_{nt}]| \\
&\quad + |EDg(F(S_{nt}))DF(S_{nt})[\Delta_n]e(W_{nt})| \\
&= I_5 + I_6 + I_7, \quad \text{say,}
\end{aligned}
\tag{2.25}
$$

where $T_{nt} = (d/dt)S_{nt}$ and t is uniformly distributed in $[0, 1]$.

Estimation of the term I_6 of (2.25). Notice that by definition of $e(x)$ as a conditional density

$$E\big(De(W_{nt})[T_{nt}]|Z_j,\ j \ne n + r\big) = 0 \quad \text{a.s. for } r = 1, \dots, B.$$

Hence

$$
\begin{aligned}
I_6 &\le E|g(F(S_{nt})) - g(F(W_{nt}))||De(W_{nt})[T_{nt}]| \\
&\le cEI\big(|F(S_{nt}) - z| \le |F(S_{nt}) - F(W_{nt})|\big)|De(W_{nt})[T_{nt}]|
\end{aligned}
\tag{2.26}
$$

by arguments similar to those used in (2.13). Using (2.22) the estimation of (2.26) is very similar to the estimations of (2.14)–(2.23), when we condition on t, split $e(x)$ with $b = B - 3$, and apply Lemma 3.12 together with Remark 3.24. Here again the splitting can be chosen such that at most third-order powers of $\|X_{n+j}\|$ occur in a final estimate similar to (2.23) together with factors of the type $\sigma(W_{(n+1)t})^{-6}$. The result is again

$$I_6 = O(n^{-1/2}) \tag{2.27}$$

by Remark 3.24 with $q = 6$.

Estimation of the term I_7 in (2.25). We have by (2.37)

$$
\begin{aligned}
I_7 &\le cn^{-1/2}|Ee(W_{nt})Dg(F(S_{nt}))DF(S_{nt})[Z_{n+1}]| \\
&\le c\big|Ee(W_{nt})I\big(|F(S_{nt}) - z| \le n^{-1/2}\big)(1 + \|S_{nt}\|^p)\|Z_{n+1}\|\big|.
\end{aligned}
$$

Conditioning on Z_{n+1} we split $e(x) = e_1(x) + r_n$ with $b = B - 2$ and may proceed as in (2.13)–(2.24) by partial integration (Lemma 3.7 together with Lemmas 3.12 and 3.24), obtaining $I_7 = O(n^{-1/2})$.

Combining this result with (2.23), (2.24), and (2.26) it follows from (2.11) and (2.12) that

$$I_1 = I_5 + O(n^{-1/2}). \tag{2.28}$$

In view of (2.8) it is sufficient to show that $I_5 = O(n^{-1/2})$ in order to prove the theorem.

Estimation of the term I_5 of (2.12). Split $e(x) = e_1(x) + r_n$ as in (2.16) with $b = B - 3$. We have in a manner similar to (2.17)–(2.23) with $\delta b > 1$,

$$I_5 \le E|Dg(F(S_{nt}))DF(S_{nt})[T_{nt}]e_1(W_{nt})| + O(n^{-1/2}). \tag{2.29}$$

Define $G(x)[y] = Dg(F(x+\Delta_{nt}))DF(x+\Delta_{nt})[y]e_1(x)$, suppressing the dependence on Δ_{nt} in this definition.

Applying Lemma 3.1(ii) (which does a second-order Taylor expansion with integral remainder term with respect to the i.i.d. summands Z_j of T_{nt}) with this function $G(x)[y]$ yields with the notation used there ($U_{nt} = S_{nt} - n^{-1/2}Z_{1t}$ and $U'_{nt} = S_{nt} - (1-\tau)n^{-1/2}Z_{1t}$)

$$I_5 = EG(U_{nt})[a_{nt}] + EL_{nt}[DG(U_{nt})] + O(n^{-1/2}) \tag{2.30}$$
$$+ED^2G(U'_{nt})[Z_{1t}^2, Y_{1t}](1-\tau)n^{-1/2}.$$

Here we have

$$\begin{aligned} D^2G(x)[z^2, y] = \{&D^3g(F(\bar{x}))DF(\bar{x})[y]DF(\bar{x})[z]^2 \\ &+ 2D^2g(F(\bar{x}))D^2F(\bar{x})[y,z]DF(\bar{x})[z] \\ &+ Dg(F(\bar{x}))D^3F(\bar{x})[z^2, y]\}e_1(x) \\ &+ D^2g(F(\bar{x}))DF(\bar{x})[y]DF(\bar{x})[z]De_1(x)[z] \\ &+ Dg(F(\bar{x}))D^2F(\bar{x})[y,z]De_1(x)[z] \\ &+ Dg(F(\bar{x}))DF(\bar{x})[y]D^2e_1(x)[z^2], \end{aligned} \tag{2.31}$$

where

$$\bar{x} = x + \Delta_{nt}.$$

The expressions for $DG(x)[z, y]$ are similar.

Summarizing we have

$$\begin{aligned} &D^iG(x)[z^i, y] \\ &= \textstyle\sum'' D^pg(F(\bar{x}))M_{pr}(x, \Delta_{nt})[z^i, y]n^{-\delta k}\sigma(x)^{-2k-2r}\prod_j{}' e(x, Z_{n+j}), \end{aligned} \tag{2.32}$$

where $i, p = 0, 1, 2$, the sum Σ'' is over all integral numbers k, p, r, and $(B-k)$-tuples of numbers $n+1$ to $n+b$ such that $p+r \le i+1$, $r \ge 0$, $k \le B-3$. The product $\prod'$ extends over $B-k$ indices $j \ge 1$ of a $(B-k)$-tuple. The functions $M_{pr}(x, \bar{x})$ are of class $C^{3+\beta-(i-p+1)}$ for $i \ge 1$ (conditionally on Δ_{nt}).

The next step is to reduce $D^pg(x)$ to $g(x)$ by repeated application of Lemma 3.7. This lemma can be applied up to three times since (2.32) shows that there are at least six factors $e(Z_{n+j}, x)^2$ in every summand of (2.32) which (even after application of Lemma 3.7) always allows us to rewrite the resulting term in the form required by (3.8).

We shall demonstrate this procedure for a typical term in the sum $e_1(x)$ (or its derivative) involving D^3g [see (2.31) resp. (2.32)] and we denote this term by I. The expectations of other terms occurring in (2.31) resp. (2.32) can be treated similarly.

In Lemma 3.7 let

$$H(x, \Delta_{nt}) = M_{3r}(x, \Delta_{nt})[Z_{1t}^2, Y_{1t}]n^{-\delta k}\sigma(x)^{-2r-2k}\prod_j{}'' e(x, Z_{n+j}),$$

where the product $\prod''$ extends over $B - k - 1$ indices $2 \le j \le B$. Furthermore, we replace expectations by conditional expectations given $\mathscr{F} = \sigma(Z_{n+j}, j > 1, Z_{1t}, t, \tau)$ and identify θ with θ_1. Let $M = n - 1$ and replace Z_j by Z_{jt} as well as g by $D^2 g$ and Δ by Δ_{nt} in Lemma 3.7. The particular term I can now be written as

$$\begin{aligned} I = n^{-1/2} E\Big\{ E\Big(\big[& D^2 g(F(S'_{nt})) DF(S'_{nt})[S'_{nt}] H(S'_{nt}, \Delta_{nt}) n/(n-1) \\ & + D^2 g(F(U'_{nt})) DF(U'_{nt})[Z_{n+1}]^2 H_1(U'_{nt}, \Delta_{nt}) \\ & + D^2 g(F(S''_{nt})) DF(S''_{nt})[a_{nt}] H(S''_{nt}, \Delta_{nt}) \\ & + n^{-1/2} D^3 g(F(U'_{nt})) H_2(U'_{nt}, \Delta_{nt}) \big] \Big| \mathscr{F} \Big) \Big\}, \end{aligned} \tag{2.33}$$

where

$$\begin{aligned} S'_{nt} &= U'_{nt} + n^{-1/2}(1 - \theta_1) Z_{(n+1)t}, \\ S''_{nt} &= U'_{nt} - n^{-1/2} \theta_1 Z_{(n+1)t}, \end{aligned} \tag{2.34}$$

and H_j, $j = 1, 2$ denote functions of class $C^{1+\beta}$ involving $\sigma(x)^{-2r-2k} \times \prod^*{}_j e(x, Z_{n+j})$, where the product extends over at least $B - k - 1 > 4$ different indices $n + j > n + 1$.

Applying Lemma 3.7 to the r.h.s. of (2.33) again with $M = n$ resp. $n - 1$ and g replaced by Dg resp. $n^{-1/2} D^2 g$ yields

$$\begin{aligned} I \le n^{-1/2} \Big| E\big[& Dg(F(S'^*_{nt})) H_4(S'^*_{nt}, \Delta_{nt}) \\ & + \text{similar terms involving } S''^*_{nt}, S^*_{nt} \\ & + n^{-1/2} D^2 g(F(U'_{nt})) H_5(U'_{nt}, \Delta_{nt}) + \text{similar terms} \\ & + n^{-1} D^3 g(F(U'_{nt})) H_6(U'_{nt}, \Delta_{nt}) + \text{similar terms} \big] \Big|, \end{aligned} \tag{2.35}$$

where S'^*_{nt}, S^*_{nt} etc., are derived from the quantities in (2.34) as before by conditioning on $\theta_2 = 0, 1$.

Here $H_j(x, \Delta)$ denote differentiable functions of class C^β [see (1.3)] such that

$$\begin{aligned} |H_j(x, \Delta)| \le{} & c(1 + \|\Delta\|^q + \|x\|^q)(1 + \sigma(x)^{-8}) n^{-\delta k} \\ & \times (1 + \|Z_{n+2}\|^2 + \|Z_1\|^3 + \|Y_1\|^3) \\ & \times \sigma(x)^{-2k-2r} \prod{}^{**} e(x, Z_{n+j}), \end{aligned} \tag{2.36}$$

where $\prod^{**}$ denotes a product over at least one index $n + j$ out of $B - k - 2$. By the choice of the function g

$$\begin{aligned} & Dg(x) \le 0, \\ & n^{-i/2} |D^{i+1} g(x)| \le c |Dg(x)| \le c n^{1/2} I(|x - z| \le n^{-1/2}). \end{aligned} \tag{2.37}$$

Let $S^{**}_{nt} = S_{nt} + O(n^{-1/2})(Z_{n+1} + Z_{n+2})$ be one of the sums obtained by the partial integrations above. Then

$$I \le c E I(F(S^{**}_{nt}) - z| \le n^{-1/2}) \sum_j \|H_j\|.$$

Using (2.37) together with (2.22) and the splitting arguments of (2.16)–(2.23) we obtain (with $q = 8$)

$$I = O(n^{-1/2})$$

for the term in (2.25) involving $D^3 g$. The other terms can be treated similarly. Hence

$$I_5 = O(n^{-1/2})$$

and the theorem holds. □

3. Lemmas. Let R_j, $j = 1, \dots, n$, denote i.i.d. Gaussian random vectors with the same distribution as R.

3.1 PROPOSITION. (i) *Let g: $\mathbb{R} \to \mathbb{R}$ denote a function of class C^1, let F denote a functional on E satisfying* (1.3), *and let H: $E \to \mathbb{R}$ denote a functional of class C^1. Let $t \in [0,1]$ be uniformly distributed and independent of X_j, $j \in \mathbb{N}$. Define*

$$S_{nt} = n^{-1/2}\{Z_{1t} + \cdots + Z_{nt} + \Delta_{nt}\} \quad \textit{and} \quad T_{nt} = \frac{d}{dt} S_{nt}, \quad \textit{where}$$

$$\Delta_{nt} = [\theta_1 Z_{n+1} + \cdots + \theta_B z_{n+B}] \sin\frac{\pi t}{2}, \qquad \Delta_n = \Delta_{n0}, \tag{3.2}$$

$$Z_{jt} = Z_j \sin\frac{\pi t}{2} + R_j \cos\left(\frac{\pi t}{2}\right) \quad \textit{and} \quad Y_{jt} = \frac{d}{dt} Z_{jt}.$$

Then we have by Taylor expansion with integral remainder term

$$\begin{aligned} &Eg(F(S_n + \Delta_n))H(S_n + \Delta_n) - Eg(F(R))H(R) \\ &\quad = E\{Dg(F(S_{nt}))DF(S_{nt})[T_{nt}]H(S_{nt}) \\ &\qquad + g(F(S_{nt}))DH(S_{nt})[T_{nt}]\}. \end{aligned} \tag{3.3}$$

(ii) *Let $G(x)[v]$, $x, v \in E$ denote a real valued function of class C^2 with respect to x and which is linear and continuous in v. Let $U_{nt} = S_{nt} - n^{-1/2} Z_{1t}$. Then ($\Delta'_{nt} = d/dt\, \Delta_{nt}$)*

$$\begin{aligned} &EG(S_{nt})[T_{nt} - \Delta'_{nt}] \\ &\quad = E\bigg\{G(U_{nt})[a_{nt}] + L_{nt}(DG(U_{nt})) \\ &\qquad + \sum_{j=1}^{n} D^2 G\big(S_{nt} - \tau n^{-1/2} Z_{jt}\big)\big[Z_{jt}^2, Y_{jt}\big] \tau n^{-3/2}\bigg\}. \end{aligned} \tag{3.4}$$

Here $\tau \in [0,1]$ is uniformly distributed and independent of X_j and t and

$$a_{nt} = n^{1/2} E Z_1 \frac{\pi}{2} \cos\frac{\pi t}{2}, \quad \textit{such that } \|a_{nt}\| = O(n^{-1/2}).$$

The function L_{nt} is a continuous linear functional on the Banach space of continuous symmetric bilinear forms on E (endowed with the strong supremum-norm topology for bilinear forms) defined by

$$L_{nt}(w) = Ew[X_1, X_1]I(\|X_1\| > n^\alpha)\frac{\pi\sin(\pi t)}{4} \tag{3.5}$$

on bilinear forms w, such that $\|L_{nt}\| = O(n^{-1/2})$.

PROOF. (i) The assertion immediately follows by Taylor expansion in t around $t = 0$ with integral remainder term.

(ii) We have

$$EG(S_{nt})[T_{nt}] = \sum_{j=1}^{n} n^{-1/2}EG(S_{nt})[Y_{jt}] \tag{3.6}$$

using the identical distribution of Z_{jt} and a Taylor expansion with integral remainder term with respect to the summand of S_{nt} which depends on Y_{jt}, we obtain a three-term expansion for (3.6).

In order to evaluate the first term of this Taylor expansion note that $EY_n = 0$. Since U_{nt} and Y_{nt} are independent we have

$$EG(U_{nt})[Y_{nt}] = EG(U_{nt})[a_{nt}].$$

The upper bound for $\|a_{nt}\|$ is a consequence of

$$EZ_n = EX_n - EX_nI(\|X_n\| > n^\alpha).$$

The second expansion term of (3.6) follows from

$$\begin{aligned} Ew[Z_{jt}, Y_{jt}] &= \big(Ew[Z_j, Z_j] - Ew[R, R]\big)\frac{\pi}{2}\cos\frac{\pi t}{2}\sin\frac{\pi t}{2} \\ &= \frac{\pi\sin(\pi t)}{4}\big(Ew[X_j, X_j] - Ew[R, R]\big) + L_{nt}(w), \end{aligned}$$

where the term in brackets vanishes by assumption (1.3) and Lemma 4.6 in Götze (1981).

The third term of the Taylor expansion of (ii) is

$$n^{-3/2}\sum_{j=1}^{n} ED^2G\big(S_{nt} - \tau n^{-1/2}Z_{jt}\big)\big[Z_{jt}^2, Y_{jt}\big]\tau/2,$$

thus proving the assertion (3.2). □

The following lemma provides the "partial integration" tool.

3.7 LEMMA. *Let $S_M = n^{-1/2}(Z_1 + \cdots + Z_m)$, $M \le n$, where Z_j, $j \in \mathbb{N}$, are i.i.d. and $E\|Z_1\|^3 < \infty$. Let g: $\mathbb{R} \to \mathbb{R}$ denote a function of class C^∞ and let H: $E \times E \to \mathbb{R}$ be of class C^p. Denote by Δ a linear combination of Z_{n+j}, $2 \le j \le B$, with coefficients of order $O(n^{-1/2})$ and let $S_{M+\tau} = \tau Z_{M+1}n^{-1/2} + S_M$,* *where* τ *is uniformly distributed in* $[0,1]$ *and independent of* Z_j, $j \in \mathbb{N}$. *Define* $a_j = EZ_1 jn^{-1/2}$ *and* $F_\Delta(s) = F(\Delta + s)$. *Then*

$$(3.8)\quad \begin{aligned} &EDg(F_\Delta(S_{M+\tau}))DF(S_M)[Z_{M+1}]^2 H(S_M, \Delta) \\ &\quad = Eg(F_\Delta(S_{M+1}))DF_\Delta(S_{M+1})[S_{M+1}]H(S_{M+1}, \Delta)nM^{-1} \\ &\qquad + Eg(F_\Delta(S_{M+\tau}))H_1(S_{M+\tau}, \Delta)[Z_{M+1}^2] \\ &\qquad + Eg(F_\Delta(S_M))H_3(S_M, \Delta)[a_M] \\ &\qquad + n^{-1/2}EDg(F_\Delta(S_{M+\tau}))H_2(S_{M+\tau\mu}, \Delta)[Z_{M+1}^3], \end{aligned}$$

where the functions H_j *are of class* C^{p-1} *and given by*

$$\begin{aligned} H_1(s, \Delta)[z^2] &= -D_s(DF_\Delta(s)[z]H(s, \Delta))[z], \\ H_2(s, \Delta)[z^3] &= -D_s(DF_\Delta(s))[z]^2 H(s, \Delta))[z] \\ &\quad - D_s\big(DF_{\Delta\nu}(s)[z]^2[\Delta n^{1/2}]H(s, \Delta)\big)[z], \\ H_3(s)[a] &= DF_\Delta(s)[a]H(s, \Delta), \end{aligned}$$

where ν, μ, *and* τ *have the uniform distribution in* $[0,1]$, *independent of all other r.v.'s.*

PROOF. Let J denote the first term on the r.h.s. of (3.8). By the i.i.d. assumption and Taylor expansion with integral remainder term we have

$$(3.9)\quad \begin{aligned} Mn^{-1}J &= Eg(F_\Delta(S_{M+1}))DF_\Delta(S_{M+1})[Z_{M+1}]Mn^{-1/2}H(S_{M+1}, \Delta) \\ &= Eg(F_\Delta(S_M))DF_\Delta(S_M)[a_M]H(S_M, \Delta) \\ &\quad + Mn^{-1}E\big\{Dg(F_\Delta(S_{M+\tau}))DF_\Delta(S_{M+\tau})[Z_{M+1}]^2 H(S_{M+\tau}, \Delta) \\ &\qquad + g(F_\Delta(S_{M+\tau}))D(DF_\Delta(S_{M+\tau})[Z_{M+1}]H(S_{M+\tau}, \Delta))[Z_{M+1}]\big\}. \end{aligned}$$

Using an additional Taylor expansion with integral remainder term of the second term of the r.h.s. of (3.9) around $\Delta = 0$ and $\tau = 0$ yields the l.h.s. of (3.8) as well as the term involving $H_2(s, \Delta)$ on the r.h.s. of (3.8). Hence the lemma is proved by (3.9). □

Let V denote a Gaussian random vector independent of R with the same distribution. Similar to Lemma 3.7 we have

3.10 LEMMA. *Let* H: $E \to \mathbb{R}$ *denote a function of class* C^p. *Then*

$$(3.11)\quad \begin{aligned} &EDg(F(R))DF(R)[V]^2 H(R) \\ &\quad = Eg(F(R))\{DF(R)[R]H(R) - D^2F(R)[V^2]H(R) \\ &\qquad - DF(R)[V]DH(R)[V]\}. \end{aligned}$$

PROOF. Since R and $(R_1 + \cdots + R_m)m^{-1/2}$ (where R_j denote independent copies of R) have the same distribution the arguments of the proof of Lemma 3.7

immediately entail (3.11) as $m \to \infty$. (Note that arbitrarily high moments of $\|R\|$ exist by Lemma 3.12.) □

3.12 LEMMA. *Let Z_j denote the random vector which equals X_j if $\|X_j\| \le n^\alpha$, where $\alpha \le \frac{1}{2}$ and is zero otherwise. Assume that X_j satisfy* (1.1)–(1.2). *Let $S_n = n^{-1/2}(Z_1 + \cdots + Z_n)$. Then*

(i) $$\sup\{E\|S_n\|^p: n \in \mathbb{N}\} < \infty,$$

(ii) $$E\|R\|^p < \infty$$

for every $p \in \mathbb{N}$.

PROOF. Chebyshev's inequality yields

$$\begin{aligned} EZ_1 &= EX_1 - EX_1 I(\|X_1\| > n^\alpha) \\ &= E\|X_1\|^3 o(n^{-(2+\gamma)\alpha}). \end{aligned}$$

Hence

$$\|ES_n\| = o(n^{-(2+\gamma)\alpha + 1/2}).$$

Using

$$\|S_n\|^p \le 2^{p-1}(\|ES_n\|^p + \|S_n - ES_n\|^p),$$

we may assume $EZ_j = 0$. By assumption the CLT holds for X_j. Hence, the results of de Acosta and Giné (1979) imply that $E\|X_1 + \cdots + X_n\|^2/n$ is uniformly bounded in n. Hence the same holds true for Z_j. Assertion (i) now follows immediately by an explicit estimate of Yurinskii (1976), page 478, (2.8). Assertion (ii) is a well-known result of Fernique (1970). □

3.13 LEMMA. *Let R denote a Gaussian random vector in E as in Lemma* 3.10. *Then*

$$P(|F(R) - z| \le \varepsilon) \le c\varepsilon. \tag{3.14}$$

PROOF. Let

$$g(x) = \int_{-\infty}^{x} I(-\varepsilon \le a - z \le \varepsilon)\, da.$$

Then $0 \le g(x) \le 2\varepsilon$. Writing

$$P(|F(R) - z| \le \varepsilon) = EI(-\varepsilon \le F(R) - z \le \varepsilon) DF(R)[V]^2/v(R),$$

the r.h.s. of (3.14) can be bounded after application of Lemma 3.10 [with $H(R) = v(R)^{-1}$] by

$$\begin{aligned} &Eg(F(R))\big\{DF(R)[R]v(R)^{-1} - D^2F(R)[V^2]v(R)^{-1} \\ &\qquad - DF(R)[V]D(\dot{v}(R)^{-1})[V]\big\} \\ &\le \varepsilon E\big\{v(R)^{-1}(1 + \|R\|^p)\|R\| + (1 + \|R\|^p)^2 v(R)^{-2}\big\} \\ &\le c\varepsilon E^q v(R)^{-2-s}, \qquad \text{for some } q > 0, 1 > s > 0 \\ &\le c\varepsilon \end{aligned}$$

by (3.23) and the choice of k_β, thus proving the assertion. □

3.15 **Lemma.** *Let* $Z_{jt} = Z_j \sin(\pi t/2) + R_j \cos(\pi t/2)$ *and let* S_{nt} *denote the sum* $n^{-1/2}(Z_{1t} + \cdots + Z_{nt})$. *Let* $Y_{jt} = d/dt Z_{jt}$. *As in* (2.11) *let* $\sigma(x)^2 = n^{-S} + E(DF(x)[Z_1])^2$. *Then*

$$\text{(i)}\quad \sup\left\{E\sigma(R + \eta_1 X_1 + \cdots + \eta_r X_r)^{-q}\|X_1\|^3 \cdots \|X_r\|^3 : n \in \mathbb{N}, |\eta_i| \le n^{-1/2}, 1 \le i \le r\right\} < D < \infty,$$

where $r = [q/2] + 1$ *and* $\delta < \min(\beta, 1/4)/4$ *implies*

$$\sup\left\{E\sigma(S_{ns})^{-q} : n \in \mathbb{N},\ s \in [0,1]\right\} \le C(D, \beta) < \infty. \tag{3.16}$$

When $E\|X_1\|^{3+\gamma} < \infty$, $\gamma > 0$ *and* $\delta < (\beta \wedge 1/4)/4 \wedge (\gamma/(6 + 2\gamma))$ *the condition*

$$\text{(ii)}\quad \sup_n E\sigma(R)^{-q} < c(1 + D) < \infty$$

is sufficient for condition (i) *to hold. Conditions* (i) *and* (ii) *hold assuming that the variance condition* (1.4) *holds with* $\gamma = 0$ *and* $\gamma > 0$, *respectively, when* $q < 2k_\beta$.

Proof. Let $S'_{ns} = S_{ns} - n^{-1/2}Z_{ns}$. Similar to relations (3.2)–(3.3) we have by Taylor expansion with integral remainder term

$$\begin{aligned} &E\sigma(S_{nt})^{-q} - E\sigma(R)^{-q} \\ &\quad = \int_t^1 ds \left\{ED\sigma(S'_{ns})^{-q}[Z_{ns}]n^{1/2} + ED^2\sigma(S'_{ns})^{-q}[Z_{ns}, Y_{ns}]\right. \\ &\qquad + ED^2\sigma(S_{ns} - n^{-1/2}\theta Z_{ns})^{-q}[Z_{ns}, Y_{ns}] \\ &\qquad \left. - ED^2\sigma(S'_{ns})^{-q}[Z_{ns}, Y_{ns}]\right\}, \end{aligned} \tag{3.17}$$

where θ is uniformly distributed in $[0,1]$. Using the relations in the proof of Lemma 3.12 the first term on the r.h.s. of (3.17) can be estimated for fixed s by

$$\begin{aligned} &cE(1 + \|S'_{ns}\|^{2p})\sigma(S'_{ns})^{-q-2}\|Z_n\|^3 n^{-(2+\gamma)\alpha + 1/2} \\ &\quad \le CE^\nu\sigma(S'_{ns})^{-(q+2)/\nu}E\|Z_1\|^3 n^{-(2+\gamma)\alpha+1/2}, \end{aligned} \tag{3.18}$$

for some $0 < \nu < 1$ arbitrarily close to 1 and $\alpha = 3/(6 + 2\gamma)$, $\gamma \ge 0$. Similar to (3.4) we have again by independence and the definition of Z_{ns}, Y_{ns}, and

$$\begin{aligned} Ew[Z_{ns}, Y_{ns}] &= \frac{\pi}{2}\cos\frac{\pi s}{2}\sin\frac{\pi s}{2}\{Ew[R, R] - Ew[X_n, X_n]I(\|X_n\| < n^\alpha)\} \\ &= O(n^{-\alpha(1+\gamma)})\|w\|E\|X_1\|^3 I(\|X_1\| > n^\alpha). \end{aligned}$$

Using this relation in order to estimate the second term of the r.h.s. of (3.17) similar as the first one we get as an upper bound for the second term

$$cE(1 + \|S'_{ns}\|^{4p})\sigma(S'_{ns})^{-(q+4)}O(n^{-\alpha(1+\gamma)})E\|X_2\|^{3+\gamma}. \tag{3.19}$$

Finally, the third term of the r.h.s. needs more elaborate estimation using

$$
\begin{aligned}
|D^2\sigma(x)^{-q} - D^2\sigma(y)^{-q}| \le c(1 + \|x\|^p + \|y\|^p)^4 \\
\times\Big\{\|x-y\|^\beta\sigma(x)^{-q-2} + \|y-y\|\sigma(x)^{-q-4} + \|x-y\| \\
\times \int d\theta\,\sigma(\theta x + (1-\theta)y)^{-q-6}(1 + \|x\|^p + \|y\|^p)^2\Big\}
\end{aligned} \tag{3.20}
$$

for every $x, y \in E$ with $\|x - y\| < 1$ and integration over $[0,1]$. This inequality follows from condition (1.3) and elementary calculus.

Relations (3.17)–(3.20) together with Hölder's inequality, $\|Z_{1s}\| < n^\alpha + \|Y_1\|$ and $\sup \sigma(x)^{-2} \le n^\delta$ yield for some $\nu < 1$ sufficiently close to 1,

$$
\begin{aligned}
E\sigma(S_{nt})^{-q} &\le E\sigma(R)^{-q} \\
&+ \int_t^1 ds\Big\{O(n^{-\alpha(1)} + n^{-\alpha(2)} + n^{-\alpha(3)} + n^{-\alpha(4)})E^\nu\sigma(S'_{ns})^{-q/\nu} \\
&+ O(n^{-\alpha(5)})\int_0^1 d\tau\, E(1 + \|S'_{ns}\|^{6p})(\|Z_n\|^3 + \|Y_n\|^3) \\
&\times \sigma(S'_{ns} + \tau n^{-1/2}Z_{ns})^{-q}\Big\},
\end{aligned} \tag{3.21}
$$

where $\alpha(1) = (2+\gamma)\alpha - 1/2 - \delta$, $\alpha(2) = (1+\gamma)\alpha - 2\delta$, $\alpha(3) = \beta/2 - \delta$, $\alpha(4) = 1/2 - 2\delta$, and $\alpha(5) = 1/2 - 3\delta$. Using

$$
E^\nu\sigma(S'_{ns})^{-q/\nu} \le n^{(1-\nu)q\delta/2}E\sigma(S'_{ns})^{-q}
$$

and $E\|Y_n\|^p < \infty$ for arbitrarily large p by Lemma 3.12(ii) we have

$$
\begin{aligned}
E\sigma(S_{nt})^{-q} &\le E\sigma(R)^{-q} + O(n^{-\chi})\int_t^1 ds\Big\{E\sigma(S'_{ns})^{-q} \\
&+ \int d\tau\, E\sigma(S'_{ns} + \tau n^{-1/2}Z_{ns})^{-q}\|Z_n\|^3\Big\} \\
&\le E\sigma(R)^{-q} + O(n^{-\chi}) \sup_{0\le\tau,\, s\le 1} E\sigma(S'_{ns} + \tau n^{-1/2}Z_{ns})^{-q}\|Z_n\|^3,
\end{aligned} \tag{3.22}
$$

where $\chi = \min(\alpha(1), \dots, \alpha(5)) - (1-\nu)q\delta/2$. Choosing ν sufficiently close to 1 and $\delta < (\beta \wedge 1/4)/4$ we obtain $\chi > (\beta \wedge 1/4)/4 > \delta$.

Conditioning on Z_n we apply (3.22) r times recursively, obtaining

$$
\begin{aligned}
E\sigma(S_{nt})^{-q} \le \sup\Big\{\sum_{j=0}^{r-1} O(n^{-\chi j})E\sigma(R\chi_{nj} + \Delta_{nj})^{-q}\|Z_n\|^3 \cdots \|Z_{n-j+1}\|^3 \\
+ O(n^{-\chi r})\sup_s E\sigma(S'_{ns} + \Delta_{nr})^{-q}\|Z_n\|^3 \cdots \|Z_{n-r}\|^3\Big\},
\end{aligned}
$$

where $\Delta_{nj} = \sum_{k=0}^j \tau_k Z_{ns_k} n^{-1/2}$, $\chi_{nj}^2 = 1 - j/n$, and the supremum is over all $0 \le \tau_k$, $s_k \le 1$. Since

$$
\sigma^2(R\chi_{nj} + \Delta_{nj}) \ge \sigma^2(R + \Delta_{nj})/2,
$$

and

$$c\left(1 + \|R\|^p + \|\Delta_{nj}\|^p\right)^2 (1 - \chi_{nj})\|R\| < n^{-\delta}/2$$

hold with probability $1 - o(n^{-A})$, $A > 0$ arbitrarily large, we have by the assumptions of Lemma 3.15,

$$\sup_{t,n} E\sigma(S_{nt})^{-q} \le cD + O(n^{-A}) + O(n^{-\chi r + q\delta/2})\left(E\|Z_n\|^3\right)^r.$$

Choosing $r = [q/2] + 1$, the proof of the first part of Lemma 3.15 is complete.

When $\gamma > 0$ we have

$$\sigma^2(R + \Delta_{nj}) \ge \sigma^2(R) - c\left(1 + \|R\|^p + \|\Delta_{nj}\|^p\right)^2 \|\Delta_{nj}\|.$$

Since $\|\Delta_{nj}\| \le cn^{\alpha - 1/2}$ we have $c(1 + \|R\|^p)^2\|\Delta_{nj}\| \ge \frac{1}{2}n^{-\delta}$ with probability $1 - O(n^{-A})$, $A > 0$ arbitrarily large, provided $\alpha - 1/2 < -\delta$, which follows from $\delta < \gamma/(6 + 2\gamma)$. Hence $\sigma^2(R + \Delta_{nj}) \ge \frac{1}{2}\sigma^2(R)$ with probability $1 - O(n^{-A})$, which immediately shows that

$$E\sigma(R)^{-q} < cD\left(E\|X_1\|^3\right)^{-r} + c < \infty$$

is a sufficient condition.

In order to reduce (i) and (ii) to the variance condition (1.4) note that by Chebyshev's inequality

$$\sigma(R)^2 = v(R) + n^{-\delta} + \left(1 + \|R\|^{2p}\right)O(n^{-(1+\gamma)\alpha})E\|X_1\|^{3+\gamma},$$

i.e.,

$$E\sigma(R)^{-k} \le Ev(R)^{-k/2} + O(n^{k\delta/2})P\left(\|R\| > n^{(\alpha-\delta)/(2p)}\right).$$

By Lemma 3.9 and condition (1.4) we have for $k > 0$, $\gamma > 0$,

$$\begin{aligned} E\sigma(R)^{-k} &\le c\int_0^1 x^{-k-1}P\left(v(R) \le x^2\right) dx + O(1) \\ &\le c\int_0^1 x^{-k-1}x^{2k_\beta}\, dx + O(1) < \infty \end{aligned} \tag{3.23}$$

for $k < 2k_\beta$. For $\gamma = 0$ the proof is similar. This completes the proof of Lemma 3.15. □

3.24 Remark. Using the notation of Lemma 3.15 we have for arbitrary $A > 0$ and $q = k_\beta/(1 + \varepsilon)$, for some $\varepsilon > 0$ sufficiently small

$$\sup\left\{E\left(1 + \|S_{nt}\|^A\right)\sigma(S_{nt})^{-q}\|Z_n\|^3 : n \in \mathbb{N},\ t \in [0,1]\right\} < \infty, \tag{3.25}$$

provided condition (1.4) holds.

Proof. Since $\|Z_n\| \le n^\alpha$ we have $\|S_{nt}\|^A \le 2^{A-1}(1 + \|S_{nt} - n^{-1/2}Z_{nt}\|^A)$. Conditioning on Z_n and Y_n we may proceed as in Lemma 3.15 using Hölder's inequality to get rid of the factor $(1 + \|S_{nt} - n^{-1/2}Z_{nt}\|^A)$, thereby replacing q by $q(1 + \varepsilon)$, F by $F(\cdot + n^{-1/2}Z_n)$, and n by $n - 1$. Using condition (1.4) we can carry out the proof of Lemma 3.15 with the additional factor $\|Z_n\|^3$ for $\gamma = 0$ and $\gamma > 0$ as well.

REFERENCES

DE ACOSTA, A. and GINÉ, E. (1979). Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. verw. Gebiete* **48** 213–231.

FERNIQUE, M. X. (1970). Intégrabilité des vecteurs Gaussiens. *C. R. Acad. Sci. Paris Sér. A* **270** 1698–1699.

GÖTZE, F. (1979). Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. verw. Gebiete* **50** 333–355.

GÖTZE, F. (1981). On Edgeworth expansions in Banach spaces. *Ann. Probab.* **9** 852–859.

GÖTZE, F. (1984). Expansions for von Mises functionals. *Z. Wahrsch. verw. Gebiete* **65** 599–625.

GÖTZE, F. (1985). Asymptotic expansions in functional limit theorems. *J. Multivariate Anal.* **16** 1–20.

HOFFMANN-JØRGENSEN, J. and PISIER, G. (1976). The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.* **4** 587–599.

KANDELAKI, N. P. (1965). On a limit theorem in Hilbert space. *News of the Computing Centre of the Academy of Sci. Georgian SSSR* **1** 46–55. (Russian).

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent rv's and the sample df, I. *Z. Wahrsch. verw. Gebiete* **32** 111–131.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent rv's and the sample df, II. *Z. Wahrsch. verw. Gebiete* **34**, 33–58.

KUELBS, J. and KURTZ, T. (1974). Berry-Esseen estimates in Hilbert space and an application to the law of the iterated logarithm. *Ann. Probab.* **2** 1–35.

NAGAEV, S. V. (1985). Speed of convergence to the normal law in Hilbert space. *Teor. Verojatnost. i Primenen.* **30** 19–32. (Russian).

NAGAEV, S. V. and CHEBOTAREV, V. I. (1978). On an estimate of the speed of convergence in the central limit theorem for vectors with values in the space l_2. In *Mathematical Analysis and Contiguous Questions of Mathematics.* Nauka, Moscow (Russian).

OPSIPOV. L. V. and ROTAR, V. I. (1984). On an infinite dimensional central limit theorem. *Theory Probab. Appl.* **29** 375–383.

PAULAUSKAS, V. I. (1976). On the rate of convergence in the central limit theorem in certain Banach spaces. *Theory Probab. Appl.* **21** 754–769.

PAULAUSKAS, V. I. (1979). The rates of convergence in the central limit theorem in Banach spaces. *Lecture Notes in Math.* **828** 234–243. Springer. Berlin.

RACHKAUSKAS, A. (1980). Asymptotic analysis of probability measures in Banach spaces. *Lithuanian Math. J.* **20** 249–253.

RHEE, W. S. and TALAGRAND, M. (1984). Bad rates of convergence for the central limit theorem in Hilbert space. *Ann. Probab.* **12** 843–850.

SAZONOV, V. V. (1981). *Normal Approximation. Some Recent Advances. Lecture Notes in Math.* **879**. Springer, Berlin.

SAZONOV, V. V. and ZALESSKII, B. A. (1985). On the central limit theorem in Hilbert space. In *Multivariate Analysis VI* (P. R. Krishnaiah, ed.) 495–526.

YURINSKII, V. V. (1976). Exponential inequalities for sums of random vectors. *J. Multivariate Anal.* **6** 473–499.

YURINSKII, V. V. (1982). On the accuracy of normal approximation in a Hilbert space. *Theory Probab. Appl.* **27** 280–289.

ZALESSKII, B. A. (1982). Estimates of the accuracy of normal approximations in a Hilbert space. *Theory Probab. Appl.* **27** 290–298.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
UNIVERSITÄTSSTRASSE
4800 BIELEFELD 1
WEST GERMANY