ON RATES OF CONVERGENCE FOR THE INVARIANCE PRINCIPLE[1]

BY

WALTER A. ROSENKRANTZ

I. Introduction. Let $(X_i)_{i=1}^{\infty}$ denote a sequence of independent and identically distributed random variables satisfying the following condition:

$$(1) \qquad E(X_i) = 0, \quad V(X_i) = \sigma^2,$$

$$(2) \qquad E(|X_i|^{2+a}) = b, \quad 0 < b < +\infty, \quad 0 < a \leq 2.$$

For each integer n we define X_{ni}, $i = 1, 2, \ldots, n$ as follows:

$$(3) \qquad X_{ni} = X_i/\sigma n^{1/2}.$$

It follows at once from (1)–(3) that the random variables, X_{ni} have the following properties:

$$(4) \qquad E(X_{ni}) = 0, \quad V(X_{ni}) = n^{-1},$$

$$(5) \qquad n^r E(|X_{ni}|^{2r}) = d < +\infty, \quad d = b\sigma^{-(2+a)} \quad \text{and} \quad 2r = 2+a.$$

Let $s_{nk} = \sum_{i=1}^{k} X_{ni}$ and denote by $W(t)$, the one-dimensional Brownian motion process with covariance function $r(s, t) = \min(s, t)$ and $0 \leq t < +\infty$. We define probability distribution functions $F_n(\lambda)$ and $F(\lambda)$ as follows:

$$(6) \qquad \begin{aligned} F_n(\lambda) &= \Pr\left\{ \operatorname*{Max}_{1 \leq k \leq n} |s_{nk}| \leq \lambda \right\}, \\ F(\lambda) &= \Pr\left\{ \operatorname*{Max}_{0 \leq t \leq 1} |W(t)| \leq \lambda \right\}. \end{aligned}$$

In their now classical paper [4], Kac and Erdös proved that $\lim_{n \to \infty} F_n(\lambda) = F(\lambda)$ where

$$(7) \qquad F(\lambda) = \frac{4}{\pi} \sum_{i=0}^{\infty} (2i+1)^{-1} \exp\left[-\frac{(2i+1)^2 \pi^2}{8\lambda^2} \right].$$

These results were later generalized by Donsker [2] and Prochorov [8] and the method employed to obtain these limit theorems has become known as "the invariance principle." The purpose of this paper is to study the rate of convergence of F_n to F. In particular we shall prove

THEOREM 1.

$$(8) \qquad |F_n(\lambda) - F(\lambda)| \leq A(\log n)^{1/2} n^{-\mu}, \quad \text{where } \mu = \tfrac{1}{2} a/(a+3)$$

and A is a constant depending on σ and b, but independent of λ.

For the case $a = 1$, i.e., $E(|X_i|^3) < +\infty$, Theorem 1 predicts

$$|F_n(\lambda) - F(\lambda)| \leq A(\log n)^{1/2}/n^{1/8}.$$

This case has also been studied by K. L. Chung [1] and Y. Prochorov [8, Theorem 4.1] who obtained respectively the following results:

THEOREM 2.1 (CHUNG).

(9) $$|F_n(\lambda) - F(\lambda)| \leq A'(\log \log n^{1/2}/\log n^{1/2})^{1/2}.$$

THEOREM 2.2 (PROCHOROV).

(10) $$|F_n(\lambda) - F(\lambda)| \leq A''(\log n)^2/n^{1/8}.$$

It is to be observed that the method used in this paper, which is based on a beautiful representation theorem of Skorokhod [9], is quite different from that of the above-mentioned authors—it is also much simpler. Next, a word about the organization of this paper: In Part II we state the Skorokhod representation theorem and apply it to the random variables defined at (3); the proof of Theorem 1 is given in Part III and an extension of Theorem 1 is given in Part IV. Finally it should be mentioned that it is still an open question what are the best possible rates of convergence. The author conjectures that $(\log n)^{1/2}/n^{1/2}$ is best possible.

II. **The Skorokhod representation.** In this section we present, without proofs, some important results due to Skorokhod. It is these results that we need to prove Theorem 1.

Let ξ be a random variable defined over a probability space (Ω_1, B_1, P_1) and such that $E(\xi) = 0$. Let $\Omega_2 = C[0, \infty)$ be endowed with the Wiener measure P_W and we denote a Brownian motion path by $W(t)$, $0 \leq t < +\infty$. We now form the product space $\Omega = \Omega_1 \times \Omega_2$ and give it the product measure $P_1 \times P_W = P$. We can now assume that ξ and $W(t)$ are random variables defined over (Ω, P) and we note that ξ and $W(t)$ are *independent* relative to P. With these preliminaries out of the way we can now state the first part of the Skorokhod representation theorem (Chapter 7 of [9]):

THEOREM 3. *There exists a nonnegative random variable* $\tau(\omega)$ *which is a Markov time relative to the Brownian motion process* $W(t, \omega)$ $(\omega \in \Omega)$, *with the following properties: The random variable* $W(\tau(\omega), \omega)$, *which we shall abbreviate to* $W(\tau)$, *has the same distribution as* ξ; *moreover the following moment inequalities are valid:*

(11) $$E(\tau) = V(\xi)$$

and more generally

(12) $$E(\tau^k) \leq C_k E(|\xi|^{2k}), \qquad k = 1, 2, \ldots,$$

where C_k *is an absolute constant independent of the distribution of* ξ.

REMARKS. The fact that τ is a Markov time for the Brownian motion process is important and is used in the following way: Let Y_1 and Y_2 be mutually independent random variables and such that $E(Y_1) = E(Y_2) = 0$. According to Skorokhod's

Theorem there exists a Markov time τ_1 with the property that Y_1 and $W(\tau_1)$ are identically distributed. Consider the process $W_1(t) = W(\tau_1 + t) - W(\tau_1)$, $0 \leqq t < \infty$; by the strong Markov property (see [6], [7]) $W_1(t)$ is independent of both τ_1 and $W(\tau_1)$. Now apply the Skorokhod representation to Y_2 and $W_1(t)$ to conclude there exists a Markov time τ_2, relative to the process $W_1(t)$, such that $W_1(\tau_2)$ and Y_2 are identically distributed. It is clear that $W_1(\tau_2)$ is independent of $W(\tau_1)$ and that τ_1 is also independent of τ_2. Thus, the joint distribution of Y_1, Y_2 is the same as the joint distribution of $W(\tau_1)$, $W_1(\tau_2) = W(\tau_1 + \tau_2) - W(\tau_1)$. This naturally leads to the following general version of the Skorokhod representation theorem.

THEOREM 4. *Let* $Y_1, Y_2, \ldots, Y_n$ *be mutually independent random variables with zero means and* $V(Y_i) = \sigma_i^2$. *Then there exists a sequence of nonnegative, mutually independent random variables* $\tau_1, \tau_2, \ldots, \tau_n$ *with the following properties.*

The joint distributions of the random variables $Y_1, Y_2, \ldots, Y_n$ are identical to the joint distributions of the random variables

$$(13) \quad W(\tau_1), \; W(\tau_1 + \tau_2) - W(\tau_1), \ldots, W(\tau_1 + \cdots + \tau_n) - W(\tau_1 + \cdots + \tau_{n-1}).$$

Moreover, we have

$$(14) \qquad\qquad E(\tau_i) = \sigma_i^2$$

and

$$(15) \qquad\qquad E(\tau_i^k) \leqq C_k E(|Y_i|^{2k}), \qquad k = 1, 2, \ldots.$$

We now apply Theorem 4 to the random variables X_{ni}, $i = 1, 2, \ldots, n$, defined at (3). Thus we can assert that there exist mutually independent, identically distributed and nonnegative random variables $\tau_{n1}, \tau_{n2}, \ldots, \tau_{nn}$ with the following properties:

$$(16) \qquad\qquad E(\tau_{ni}) = n^{-1},$$

and in addition the random variables

$$(17) \quad W(\tau_{n1}), \; W(\tau_{n1} + \tau_{n2}) - W(\tau_{n1}), \ldots, W(\tau_{n1} + \cdots + \tau_{nn}) - W(\tau_{n1} + \cdots + \tau_{nn-1})$$

have the same joint distribution as the random variables X_{ni}, $i = 1, 2, \ldots, n$. In particular the joint distribution of $s_{n1}, s_{n2}, \ldots, s_{nn}$ is the same as the joint distribution of the random variables:

$$(18) \qquad\qquad W(\tau_{n1}), \; W(\tau_{n1} + \tau_{n2}), \ldots, W(\tau_{n1} + \cdots + \tau_{nn}).$$

We now define random variables z_{nk}, $k = 1, 2, \ldots, n$ as follows:

$$(19) \qquad\qquad z_{nk} = \sum_{i=1}^{k} (\tau_{ni} - n^{-1}).$$

Thus (18) can be rewritten in the following more useful form:

$$(20) \qquad\qquad W(1/n + z_{n1}), \; W(2/n + z_{n2}), \ldots, W(1 + z_{nn}).$$

Since z_{nk}, $k=1, 2, \ldots, n$ is a sequence of partial sums of mutually independent, identically distributed random variables, each with zero means, it is reasonable to expect that if $E(|\tau_{ni} - n^{-1}|^{\gamma}) < +\infty$ for some $\gamma > 1$ then we can obtain an estimate on the rate at which $\Pr\{\text{Max}_{1 \leq k \leq n} |z_{nk}| > \delta\}$ tends to zero. Kolmogorov's inequality ([7, p. 235]) is not applicable because our hypotheses are not strong enough to insure the existence of a second moment for the τ_{ni}, $i=1, 2, \ldots, n$. To get such estimates it is first of all necessary to extend inequality (15) as follows:

LEMMA 1.

(21) $$E(\tau_i^{\gamma}) \leq C_{\gamma} E(|Y_i|^{2\gamma}) \quad \text{for any } \gamma \geq 1.$$

The proof of this result is given in the appendix.

An immediate consequence of Lemma 1 (set $\gamma = r$) is the following result:

(22) $\quad v_n = E(|\tau_{ni} - n^{-1}|^r) \leq 2^{r-1}\{E(\tau_{ni}^r) + n^{-r}\} \leq 2^{r-1}C_r E(|X_{ni}|^{2r}) + 2^{r-1}n^{-r}.$

By (5) one easily concludes

(23) $$n^r v_n \leq B < +\infty,$$

where

$$B = 2^{r-1}(C_r d + 1).$$

The Kolmogorov inequality for martingales (cf. [7]) yields:

(24) $$\Pr\left\{\underset{1 \leq k \leq n}{\text{Max}} |z_{nk}| > \delta\right\} \leq \frac{E(|z_{nn}|^r)}{\delta^r}.$$

We now employ a result of von Bahr and Esseen [10] and obtain the following inequalities[2]:

(25) $$E(|z_{nn}|^r) \leq 2nv_n,$$

(26) $$n^{r-1}E(|z_{nn}|^r) \leq 2B < +\infty,$$

and thus we obtain the estimate:

(27) $$\Pr\left\{\underset{1 \leq k \leq n}{\text{Max}} |z_{nk}| > \delta\right\} \leq \frac{2B}{n^{r-1}\delta^r}.$$

Summing up, then, we have the result that the joint distributions of the random variables s_{nk}, $k=1, 2, \ldots, n$, are identical to the joint distributions of the random variables defined at (20), and so it is sufficient to prove Theorem 1 for these random variables instead of the original sequence s_{nk}, $k=1, 2, \ldots, n$.

III. **Proof of Theorem 1.** The proof uses certain facts about the "tied down Brownian motion process"; we refer the reader to the paper by Doob [3] for a more complete discussion.

---

[2] For this reference I am indebted to Professor Melvin Katz of the University of New Mexico.

Consider the following stochastic process $z_n(t)$: $z_n(t) = nt W(1/n) - W(t)$, $0 \leq t \leq 1/n$. We note that $z_n(0) = z_n(1/n) = 0$ and that $z_n(t)$ is a Gaussian process with covariance matrix $r_n(s, t) \doteq s(1 - nt)$, $0 \leq s < t \leq 1/n$. We shall refer to the $z_n(t)$ process as the Brownian motion process "tied down at 0 and n^{-1}".

LEMMA 2. $\Pr\{\text{Max}_{0 \leq s \leq n^{-1}} |z_n(s)| \geq \varepsilon\} \leq 2 \exp(-2n\varepsilon^2)$, $\varepsilon > 0$.

Proof. Following Doob [3] one observes that $(1 + nt)z_n(t/(1 + nt)) = W(t)$, $0 \leq t < +\infty$. Thus

$$\Pr\left\{\underset{0 \leq s \leq n^{-1}}{\text{Max}} |z_n(s)| \geq \varepsilon\right\} = \Pr\left\{\underset{0 \leq t < +\infty}{\text{Max}} |W(t)|/\varepsilon(1 + nt) \geq 1\right\}.$$

Now using Doob's estimates [3, pp. 397–398] we conclude that

$$\Pr\left\{\underset{0 \leq t < +\infty}{\text{Max}} |W(t)|/\varepsilon(1 + nt) \geq 1\right\} \leq 2 \exp(-2n\varepsilon^2),$$

and this completes the proof of Lemma 2.

We now introduce a probability distribution function $\hat{F}_n(\lambda)$ as follows:

$$(28) \qquad \hat{F}_n(\lambda) = \Pr\left\{\underset{1 \leq k \leq n}{\text{Max}} \left|W\left(\frac{k}{n}\right)\right| \leq \lambda\right\},$$

where $W(k/n)$ denotes Brownian motion at time $t = k/n$. We now define a "random broken line" in the sense of Prochorov [8] as follows:

$$(29) \qquad W_n(t) = W(k/n) \qquad \text{for } t = k/n,$$

and

$$W_n(t) = W\left(\frac{k}{n}\right) + n\left[W\left(\frac{k+1}{n}\right) - W\left(\frac{k}{n}\right)\right]\left(t - \frac{k}{n}\right) \qquad \text{for } \frac{k}{n} \leq t \leq \frac{k+1}{n}.$$

In other words $W_n(t)$ is a polygonal approximation to the Brownian motion path $W(t)$. Moreover it is easily verified that

$$(30) \qquad \hat{F}_n(\lambda) = \Pr\left\{\underset{0 \leq t \leq 1}{\text{Max}} |W_n(t)| \leq \lambda\right\}.$$

LEMMA 3.

$$\Pr\left\{\underset{0 \leq t \leq 1}{\text{Max}} |W(t) - W_n(t)| \geq \varepsilon_n\right\} \leq 2n \exp(-2n\varepsilon_n^2) = r_n,$$

where ε_n is a sequence tending to zero and whose precise dependence on n will be specified later.

Proof. It is only necessary to observe that $W(t) - W_n(t)$ over each of the intervals $[k/n, (k+1)/n]$, $0 \leq k \leq n-1$, is the $z_n(t)$ process defined in Lemma 2. Thus

$$\Pr\left\{\underset{0 \leq t \leq 1}{\text{Max}} |W(t) - W_n(t)| \geq \varepsilon_n\right\} \leq \sum_{k=0}^{n-1} \Pr\left\{\underset{k/n \leq t \leq (k+1)/n}{\text{Max}} |W(t) - W_n(t)| \geq \varepsilon_n\right\}$$

$$= \sum_{k=0}^{n-1} \Pr\left\{\underset{0 \leq s \leq n^{-1}}{\text{Max}} |z_n(s)| \geq \varepsilon_n\right\}$$

$$\leq 2n \exp(-2n\varepsilon_n^2).$$

An immediate consequence of Lemma 3 is

LEMMA 4.

$$F(\lambda - \varepsilon_n) - r_n \leq \hat{F}_n(\lambda) \leq F(\lambda + \varepsilon_n) + r_n.$$

Now the distribution function F is quite smooth; it even has a bounded first derivative, and thus we know there exists a constant A_1 such that

(31) $$|F(\lambda + h) - F(\lambda)| \leq A_1 |h|,$$

where A_1 is of course independent of λ. We thus conclude

LEMMA 5.

$$F(\lambda) - A_1 \varepsilon_n - r_n \leq \hat{F}_n(\lambda) \leq F(\lambda) + A_1 \varepsilon_n + r_n.$$

We now proceed to obtain a similar estimate for $\hat{F}_n$ and F_n. We remind the reader that according to the Skorokhod representation theorem:

(32) $$F_n(\lambda) = \Pr \left\{ \operatorname*{Max}_{1 \leq k \leq n} \left| W\left(\frac{k}{n} + z_{nk}\right) \right| \leq \lambda \right\}.$$

Now let B_n denote the following event:

(33) $$B_n = \left\{ \operatorname*{Max}_{1 \leq k \leq n} \left| W\left(\frac{k}{n} + z_{nk}\right) - W\left(\frac{k}{n}\right) \right| \geq \varepsilon_n \right\}$$

and note $B_n = B_{n1} \cup B_{n2}$ where

$$B_{n1} = B_n \cap \left\{ \operatorname*{Max}_{1 \leq k \leq n} |z_{nk}| \leq \delta_n \right\}$$

and

$$B_{n2} = B_n \cap \left\{ \operatorname*{Max}_{1 \leq k \leq n} |z_{nk}| > \delta_n \right\},$$

where $\delta_n > 0$ is a sequence whose dependence on n will be specified later.

LEMMA 6.

$$\Pr \{B_n\} \leq \frac{8_n}{(2\pi)^{1/2}} \frac{(\delta_n)^{1/n}}{\varepsilon_n} \exp \left(-\frac{\varepsilon_n^2}{2\delta_n^2} \right) + \frac{2B}{n^{r-1} \delta_n^r} = \rho_n.$$

Proof. We first note that

$$\Pr \{B_{n2}\} \leq \Pr \left\{ \operatorname*{Max}_{1 \leq k \leq n} |z_{nk}| > \delta_n \right\} \leq \frac{2B}{n^{r-1} \delta_n^r};$$

this last inequality is an immediate consequence of (27).

As for B_{n1} we observe that $B_{n1} \subset E_{n1} \cup E_{n2}$, where

$$E_{n1} = \bigcup_{k=1}^{n} \left\{ \operatorname*{Max}_{0 \leq s \leq \delta_n} \left| W\left(\frac{k}{n} + s\right) - W\left(\frac{k}{n}\right) \right| \geq \varepsilon_n \right\}$$

and

$$E_{n2} = \bigcup_{k=1}^{n} \left\{ \operatorname*{Max}_{0 \leq s \leq \delta_n} \left| W\left(\frac{k}{n} - s\right) - W\left(\frac{k}{n}\right) \right| \geq \varepsilon_n \right\}.$$

It is clear that $\Pr\{E_{n1}\} = \Pr\{E_{n2}\}$ and thus we conclude

$$\Pr\{B_{n1}\} \leq 2n \Pr\left\{\underset{0 \leq t \leq \delta_n}{\text{Max}} |W(t)| \geq \varepsilon_n\right\}.$$

It is easily shown that,

$$\Pr\left\{\underset{0 \leq t \leq \delta_n}{\text{Max}} |W(t)| \geq \varepsilon_n\right\} = \Pr\left\{\underset{0 \leq t \leq 1}{\text{Max}} |W(t)| \geq \frac{\varepsilon_n}{(\delta_n)^{1/2}}\right\}$$

$$\leq 4 \Pr\{W(1) \geq \varepsilon_n/(\delta_n)^{1/2}\}.$$

As is well known, $\Pr\{W(1) \geq (\varepsilon_n/(\delta_n)^{1/2})\} \leq (1/(2\pi)^{1/2})((\delta_n)^{1/2}/\varepsilon_n) \exp(-\varepsilon_n^2/2\delta_n)$ (cf. [5, Chapter 7, Vol. 1]). Putting these inequalities together we conclude:

$$\Pr\{B_{n1}\} \leq (8_n/(2\pi)^{1/2})((\delta_n)^{1/n}/\varepsilon_n) \exp(-\varepsilon_n^2/2\delta_n)$$

and this completes the proof of Lemma 6. We have therefore the following inequality:

$$(34) \qquad \hat{F}_n(\lambda - \varepsilon_n) - \rho_n \leq F_n(\lambda) \leq \hat{F}_n(\lambda + \varepsilon_n) + \rho_n.$$

Applying Lemma 5 with $\lambda \pm \varepsilon_n$ in the place of λ we conclude:

$$(35) \qquad F(\lambda - \varepsilon_n) - A_1\varepsilon_n - r_n - \rho_n \leq F_n(\lambda) \leq F(\lambda + \varepsilon_n) + A_1\varepsilon_n + r_n + \rho_n.$$

Finally by using (31) we obtain the following estimate:

$$(36) \qquad |F_n(\lambda) - F(\lambda)| \leq 2A_1\varepsilon_n + r_n + \rho_n.$$

It only remains to choose ε_n and δ_n tending to zero at just the correct rate of speed; our choices are:

$$(37) \qquad \delta_n = n^{-\alpha}, \quad \varepsilon_n = 2(\log n)^{1/2}n^{-\alpha/2}, \qquad 0 < \alpha < (r-1)/r.$$

We assume $n \geq 2$ becuase for $n=1$, $\varepsilon_1 = 0$. With these choices for δ_n and ε_n it is easily seen that $(\varepsilon_n^2/2\delta_n) = 2\log n$. Thus, for $n \geq 2$,

$$(38) \qquad \rho_n \leq \frac{4}{(2\pi)^{1/2}} \frac{n}{(\log n)^{1/2}} \cdot \frac{1}{n^2} + \frac{2B}{n^{r-1-\alpha r}} \leq \frac{2(B+1)}{n^{r-1-\alpha r}}.$$

According to Lemma 3 we have the following estimate for r_n:

$$(39) \qquad r_n = 2n \exp(-8n^{1-\alpha}\log n) \leq 2n \exp(-8\log n) = 2/n^7.$$

So r_n is negligible in comparison to ρ_n and ε_n. Thus we can assert that

$$(40) \qquad r_n + \rho_n \leq 2(B+2)/n^{r-1-\alpha r}, \qquad n = 2, 3, \ldots.$$

Hence (36) becomes:

$$(41) \qquad |F_n(\lambda) - F(\lambda)| \leq \frac{4A_1(\log n)^{1/2}}{n^{\alpha/2}} + \frac{2(B+2)}{n^{r-1-\alpha r}},$$

where A_1 and $2(B+2)$ are independent of our choice of α. The rate of the convergence is determined by $\text{Min}(\alpha/2, r-1-\alpha r)$ and so the best choice for α is obviously

Max Min $(\alpha/2, r-1-\alpha r)$, $0 < \alpha < (r-1)/r$. The solution μ satisfies the following simple linear equation: $\alpha/2 = r - 1 - \alpha r$, which means that $\alpha = (r-1)/(r+1/2)$ and thus

$$\mu = \frac{\alpha}{2} = \frac{r-1}{2r+1} = \frac{1}{2}\frac{a}{a+3}.$$

We have thus shown that we can choose α so that:

(42) $$|F_n(\lambda) - F(\lambda)| \leq A(\log n)^{1/2}/n^\mu$$

where

$$A = 4A_1 + 2(B+2), \quad \text{and} \quad \mu = \tfrac{1}{2}a/(a+3).$$

This completes the proof of Theorem 1. It is to be observed that A_1 is an absolute constant, independent of the distribution of X_i, whereas B depends on σ^2 and the $(2+a)$th moment of X_i.

IV. An extension of Theorem 1. Let $C(0; 1)$ denote the space of continuous functions on the interval $[0; 1]$ endowed with the sup norm topology i.e., if $f(t)$ and $g(t)$ are both elements of $C(0; 1)$ then we denote the distance between them in the usual way:

(43) $$\|f-g\| = \max_{0 \leq t \leq 1} |f(t) - g(t)|.$$

We denote by P_W the Wiener measure on $C(0; 1)$. By a "functional" $\mathscr{F}$ we mean a real-valued function $\mathscr{F}$ with domain $C(0; 1)$. The functional $\mathscr{F}$ is said to be *uniformly continuous* if there exists a positive constant K with the property that

(43') $$|\mathscr{F}(f) - \mathscr{F}(g)| \leq K\|f-g\|.$$

We define a "random broken line" $S_n(t)$ as follows:

(44)
$$S_n(t) = s_{nk} \qquad \text{for } t = \frac{k}{n},$$
$$S_n(t) = S_n\left(\frac{k}{n}\right) + n\left[S_n\left(\frac{k+1}{n}\right) - S_n\left(\frac{k}{n}\right)\right]\left(t - \frac{k}{n}\right), \qquad \text{for } \frac{k}{n} \leq t \leq \frac{k+1}{n}.$$

We define

(45) $$\Psi_n(\lambda) = \Pr\{-\infty < \mathscr{F}(S_n(t)) \leq \lambda\}$$

and

$$\Psi(\lambda) = P_W\{-\infty < \mathscr{F}(f) \leq \lambda\}.$$

THEOREM 5. *Let $\mathscr{F}$ be a uniformly continuous functional and suppose there exists a constant $L > 0$ with the property that*

$$|\Psi(\lambda+h) - \Psi(\lambda)| \leq L|h|.$$

Then there exists a constant $A_1 > 0$ such that

$$|\Psi_n(\lambda) - \Psi(\lambda)| \leq A_1(\log n)^{1/2}n^{-\mu}, \qquad \mu = \tfrac{1}{2}a/(a+3).$$

Proof. We first observe that

$$\Pr\{|\mathscr{F}(S_n(t)) - \mathscr{F}(W_n(t))| \geq \varepsilon_n\} \leq \Pr\{\|S_n - W_n\| \geq \varepsilon_n/K\} = \rho_n'$$

and similarly

$$\Pr\{|\mathscr{F}(W_n(t)) - \mathscr{F}(W(t))| \geq \varepsilon_n\} \leq \Pr\{\|W_n(t) - W(t)\|\} \geq \varepsilon_n/K\} = r_n'.$$

We thus obtain the following analogue of inequality (36):

(46) $$|\Psi_n'(\lambda) - \Psi(\lambda)| \leq 2L\varepsilon_n + r_n' + \rho_n'.$$

By making the same choices for ε_n and δ_n as we did in the proof of Theorem 1 we see that r_n' and ρ_n' are of the same order of magnitude as r_n and ρ_n respectively. Thus, except for a different constant, we have shown that the same order of convergence is valid for the larger class of functionals $\mathscr{F}$ satisfying the hypotheses of Theorem 5. It should also be remarked that rates of convergence can also be obtained for functionals $\mathscr{F}$ which are assumed to be merely uniformly Hölder continuous of order ν, $0 < \nu < 1$, i.e., replace (43') by the following condition[3]:

$$|\mathscr{F}(f) - \mathscr{F}(g)| \leq K\|f - g\|^\nu, \qquad 0 < \nu < 1.$$

We shall leave it to the reader to explore these and other refinements.

V. Appendix. The purpose of this part of the paper is to give a proof of inequality (21). A close examination of Skorokhod's proof of his inequality (15) shows that it suffices to establish the following inequality on the moments of the first passage times of Brownian motion. Let $a < 0 < b$ and let

$$\tau(w) = \inf_{t>0} (w(t) - a)(w(t) - b) = 0,$$

i.e., τ is the first time that the Brownian motion starting at 0 hits a or b. We shall prove the following inequality:

THEOREM 6. *There exists a constant C_r, $r \geq 1$ such that $E(\tau^r) \leq C_r|ab| \cdot |b-a|^{2r-2}$.*

In order to avoid absolute value signs we make the following simple translation of the coordinate axis: we set $x = |a|$, $y = b + |a|$, we begin the Brownian motion at x and denote by τ_x, the first time that the Brownian motion starting at x hits 0 or y. Then the inequality to be proved is the following:

(47) $$E(\tau_x^r) \leq C_r x(y - x)y^{2r-2}.$$

In the course of the proof of this result two elementary inequalities are needed, which we state now in order to avoid interrupting the argument at a later stage.

(48) $$|\sin mx| \leq m|\sin x|, \qquad m = 1, 2, \ldots.$$

---

[3] This observation is due to Dr. Stanley Sawyer of the Courant Institute of Mathematical Sciences.

Inequality (48) is trivially true for $m=1$ and it is easily proved in general by mathematical induction. Another useful inequality is the following:

$$(49) \qquad \underset{0 \leq x \leq y}{\text{Max}} \left| \frac{\sin (\pi x / y)}{x(y-x)} \right| \leq \frac{4}{y^2}, \quad 4 = \underset{0 \leq t \leq 1}{\text{Max}} \left| \frac{\sin \pi t}{t(1-t)} \right|.$$

Proof. Set $t = x/y$, holding y fixed. Then

$$\frac{\sin (\pi x / y)}{x(y-x)} = \frac{1}{y^2} \frac{\sin \pi t}{t(1-t)}.$$

Now as x varies between 0 and y, t varies between 0 and 1. Thus,

$$\underset{0 \leq x \leq y}{\text{Max}} \left| \frac{\sin (\pi x / y)}{x(y-x)} \right| = \frac{1}{y^2} \underset{0 \leq t \leq 1}{\text{Max}} \left| \frac{\sin \pi t}{t(1-t)} \right| = \frac{4}{y^2},$$

and this completes the proof of (49).

Let $H(t) = \Pr \{ \tau_x \leq t \}$. The following infinite series for $1 - H(t)$ is well known (cf. [5, Vol. 2, p. 330]):

$$(50) \qquad 1 - H(t) = \frac{4}{\pi} \sum_{n=0}^{\infty} (2n+1)^{-1} \exp \left(-\frac{(2n+1)^2 \pi^2 t}{2y^2} \right) \sin \frac{(2n+1)\pi x}{y}.$$

It follows at once from an integration by parts that

$$(51) \qquad \int_0^\infty t^r \, dH(t) = r \int_0^\infty t^{r-1}(1 - H(t)) \, dt = E(\tau_x^r).$$

We now apply inequalities (48) and (49) successively to conclude:

$$(52) \qquad \left| \sin \frac{(2n+1)\pi x}{y} \right| \leq (2n+1) \left| \sin \frac{\pi x}{y} \right| \leq 4(2n+1) \frac{x(y-x)}{y^2}.$$

Hence

$$(53) \qquad 0 \leq 1 - H(t) \leq \frac{16 x(y-x)}{\pi y^2} \sum_{n=0}^{\infty} \exp \left(-\frac{(2n+1)^2 \pi^2 t}{2y^2} \right).$$

Integrating term by term, as we may, we conclude

$$(54) \quad r \int_0^\infty t^{r-1}(1 - H(t)) \, dt \leq \frac{16 r x(y-x)}{\pi y^2} \sum_{n=0}^{\infty} \int_0^\infty t^{r-1} \exp \left(-\frac{(2n+1)^2 \pi^2 t}{2y^2} \right) dt.$$

Now a simple change of variables yields

$$(55) \qquad \int_0^\infty t^{r-1} \exp \left(-\frac{(2n+1)^2 \pi^2 t}{2y^2} \right) dt = \frac{2^r y^{2r} \Gamma(r)}{(2n+1)^{2r} \pi^{2r}},$$

where Γ denotes the Gamma function. Thus (54) becomes

$$(56) \qquad r \int_0^\infty t^{r-1}(1 - H(t)) \, dt \leq C_r x(y-x) y^{2r-2},$$

where

$$C_r = 2^{4+r}\Gamma(r+1)\pi^{-2r}\sum_{n=0}^{\infty}(2n+1)^{-2r}.$$

This completes the proof of (47) and, as we have already remarked, this result when applied to Skorokhod's proof of (15) yields inequality (21).

Acknowledgements. In the course of obtaining these results the author benefited from useful conversations with Professor S. R. S. Varadhan of the Courant Institute of Mathematical Sciences. I also wish to thank the Department of Statistics of the University of California, Berkeley, where parts of this paper were written, for the use of their facilities.

REFERENCES

1. K. L. Chung, *On the maximum partial sums of sequences of independent random variables*, Trans. Amer. Math. Soc. **64** (1948), 205–233.

2. M. Donsker, *An invariance principle for certain probability limit theorems*, Mem. Amer. Math. Soc. No. 6 (1951), 12 pp.

3. J. L. Doob, *Heuristic approach to the Kolmogorov-Smirnov theorems*, Ann. Math. Statist. **20** (1949), 393–403.

4. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. **52** (1946), 292–302.

5. W. Feller, *An introduction to probability theory and its applications*, Vols. 1 and 2, Wiley, New York, 1956; 1966.

6. G. A. Hunt, *Markoff processes and potentials*. I, Illinois J. Math. **1** (1957), 44–93.

7. M. Loève, *Probability theory*, 3rd ed., Van Nostrand, Princeton, N. J., 1963.

8. Ju. V. Prochorov, *Convergence of random processes and limit theorems in probability*, Theor. Probability Appl. **1** (1956), 157–214.

9. A. V. Skorokhod, *Studies in the theory of random processes* (Translated from Russian), Addison-Wesley, Reading, Mass., 1965.

10. B. von Bahr and C.-G. Esseen, *Inequalities for the rth absolute moment of a sum of random variables*, $1 \leq r \leq 2$, Ann. Math. Statist. **36** (1965), 299–303.

NEW YORK UNIVERSITY,
    NEW YORK, NEW YORK