The Annals of Probability
1980, Vol. 8, No. 2, 241–251

LIMIT DISTRIBUTIONS FOR THE ERROR IN APPROXIMATIONS OF STOCHASTIC INTEGRALS

BY HOLGER ROOTZÉN

University of Lund and University of Copenhagen

We consider the approximation of an Itô integral $\int_0^t \phi(s)dB(s)$ by a sequence of integrals $\int_0^t \phi_n(s)dB(s)$ of simpler integrands. It is proved that if, for a sequence $\{\psi_n\}$ of adapted integrands, $\sup_{0\leqslant t\leqslant 1}|\int_0^t \psi_n ds| \rightarrow_p 0$ and $\int_0^t \psi_n^2 ds \rightarrow_p \tau(t)$, for some continuous stochastic process $\{\tau(t); t \in [0, 1]\}$, then $\int_0^{(\cdot)}\psi_n dB \rightarrow_d W \circ \tau$ in $C(0, 1)$, where W is a Brownian motion independent of τ. Further, if one is only interested in the limit distribution of functionals like $\int_0^1\psi_n dB$ or $\sup_{0\leqslant t\leqslant 1}|\int_0^t\psi_n dB|$, then in the second condition it is enough to require that $\int_0^1\psi_n^2 ds \rightarrow_p \tau(1)$. The convergence is stable in the sense of Rényi, and from this follow results on the fluctuations of the sample paths of the integrals. As an example we consider the case $\phi(t) = f(B(t), t)$ and $\phi_n(t) = \sum_{i=1}^n f(B(i/n), i/n)I(i/n \leqslant t < (i+1)/n)$. Denoting the approximation error $\int_0^t(\phi - \phi_n)dB$ by $d_n(t)$, it follows from the above results that if f is smooth enough then $n^{\frac{1}{2}}d_n \rightarrow_d W \circ \tau$, with $\tau(t) = 2^{-1}\int_0^t f_1(B(s), s)^2 ds$ where $f_1(x, t) = \dfrac{\partial f(x, t)}{\partial x}$. Similar results are obtained for approximations of the Stratonovich integral and for higher order approximations.

0. Introduction. The Itô integral $\int_0^t\phi(s)dB(s)$ of an adapted integrand ϕ with respect to a Brownian motion process B is defined as the limit of integrals of approximating integrands ϕ_n, and if one wants to compute stochastic integrals numerically, similar approximations can be used. In this paper we investigate one aspect of the accuracy of approximations of stochastic integrals. Writing $d_n(t)$ for the approximation error $\int_0^t\phi dB - \int_0^t\phi_n dB$ we find conditions which make $a_n d_n$ converge in distribution for some suitably chosen sequence $\{a_n\}$ of normalizing constants. Furthermore it turns out that the convergence is stable in the sense of Rényi, and that this implies that the realizations $a_1d_1, a_2d_2, \ldots$ fluctuate strongly.

The main result, which will be used to prove convergence, is contained in Section 1. It is that if, for a sequence $\{\psi_n\}$ of adapted integrands, $\sup_{0\leqslant t\leqslant 1}|\int_0^t\psi_n ds| \rightarrow_p 0$ and $\tau_n(t) = \int_0^t\psi_n^2 ds \rightarrow_p \tau(t)$, $t \in [0, 1]$, for some continuous random process $\{\tau(t); t \in [0, 1]\}$, then $\int_0^{(\cdot)}\psi_n dB \rightarrow_d W \circ \tau$ in $C(0, 1)$, where W is a Brownian motion independent of τ. ($W \circ \tau$ denotes the function whose value at t is $W(\tau(t))$. In integrals dB signifies that it is an Itô integral, and thus, e.g., $\int_0^t f(B(s))dB(s)$ is often written $\int_0^t f(B)dB$, which should not be confused with the Lebesgue integral $\int_0^t f(s)ds$. Further $\int_0^t f(B)ds$ is the same as $\int_0^t f(B(s))ds$ or $\int_0^t f \circ B\, ds$.)

Further, if the second assumption is weakened to $\int_0^1\psi_n^2 ds \rightarrow_p \tau(= \tau(1))$ the one-dimensional limit results still hold, e.g., $\int_0^1\psi_n dB \rightarrow_d \int\Phi(\cdot/s^{\frac{1}{2}})dF_\tau(s)$ and

Received November 1977; revised August 1978.
AMS 1970 *subject classifications.* Primary 60H05, 60B10; secondary 60G45.
Key words and phrases. Approximations of stochastic integrals, stability in the sense of Rényi, fluctuations of approximation error.

$\sup_{0 \leq t \leq 1}|\int_0^1 \psi_n dB| \to_d \int G(\cdot / s^{\frac{1}{2}}) dF_\tau(s)$, where G is the distribution function of $\sup_{0 \leq t \leq 1}|B(t)|$. The idea of the proofs is as follows: It is known that $\int_0^{\tau_n^{-1}} \psi_n dB$ is distributed as a Brownian motion, since $\tau_n(t)$ is the "square variation" of $\int_0^t \psi_n dB$. If one can prove a little more, say that $\{\int_0^{\tau_n^{-1}} \psi_n dB\}_{n=1}^\infty$ is Rényi-mixing, then it follows that $(\int_0^{\tau_n^{-1}} \psi_n dB, \tau_n) \to_d (W, \tau)$, and thus by standard arguments that $\int_0^{(\cdot)} \psi_n dB =_{\text{a.s.}} \int_0^{\tau_n^{-1} \circ \tau_n} \psi_n dB \to_d W \circ \tau$. The proof that $\{\int_0^{\tau_n^{-1}} \psi_n dB\}$ is Rényi-mixing is accomplished by using Girsanov's theorem and a result of Duncan (1970) on absolute continuity with respect to a Brownian motion.

In Section 2 we treat approximation of integrals, with integrands of the special form $\phi(t) = f(B(s), s)$. For simplicity it is assumed that $\phi_n(t) = \sum_{i=0}^{n-1} \phi(i/n) I\{i/n \leq t < (i+1)/n\}$ and thus $\int_0^t f(B(s), s) dB(s)$ is approximated by $\int_0^t \phi_n dB = \sum_{i=0}^{[nt]-1} f(B(i/n), i/n)(B((i+1)/n) - B(i/n)) + f(B([nt]/n), [nt]/n)(B(t) - B([nt]/n))$. It is shown that if f is smooth enough then the approximation error multiplied by $n^{\frac{1}{2}}$ converges: more specifically $n^{\frac{1}{2}} d_n \to_d W \circ \tau$, where W and τ are independent and where $\tau(t) = 2^{-1} \int_0^t f_1(B(s), s)^2 ds$.

Stratonovich (1966) introduced a "symmetrized" stochastic integral. In Section 3 the accuracy of approximations of this integral is discussed using the convenient setting of Yor (1976). This setting also includes the "Rieman-Stieltjes approximations" of Wong and Zakai (1965), which are obtained by making polygonal approximations of the Brownian paths and then computing ordinary Rieman-Stieltjes integrals using these approximations. In this case the approximations are exactly equal to their limit for integrands $f(B(s))$, so the approximation error then is zero. Further, the section contains results for higher order approximations of Itô integrals. Using the methods developed in Sections 1 and 2 it is straightforward, even if somewhat cumbersome, to find limiting distributions for the errors in these approximation schemes. The structures of the limits are rather interesting. As an example can be mentioned that when the Stratonovich integral of $f(B(s))$ is approximated by sums of the form $\sum_i f(2^{-1}(B((i+1)/n) + B(i/n)))(B((i+1)/n) - B(i/n))$ then the proper normalization of the error is by n and the limiting distribution is that of $W(96^{-1} \int_0^t f''(B(s))^2 ds) + 8^{-1} \int_0^t f''(B(s)) dB(s) + 16^{-1} \int_0^t f'''(B(s)) ds$, where W is a Brownian motion independent of B.

It is possible to adapt the results of this paper to integrals with respect to more general continuous local martingales than the Brownian motion, but the next generalization—to semimartingales—does not seem to be interesting since in that case it may well be that it is the bounded variation term which determines the speed of convergence. In a further paper similar problems for approximate solutions to stochastic differential equations will be treated.

1. Limit distributions of the approximation error. Let $\{B(t); t \in [0, 1]\}$ be a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and let as usual $\{\mathcal{F}(t); t \in [0, 1]\}$ be an increasing family of σ-algebras such that $B(t) \in \mathcal{F}(t)$, $t \in [0, 1]$, and with the σ-algebra generated by $\{B(s) - B(t); t < s \leq 1\}$ independent of $\mathcal{F}(t)$ for $t \in [0, 1]$.

In subsequent sections the errors in approximations of stochastic integrals will themselves be stochastic integrals $\int_0^t a_n(\phi - \phi_n)dB$ and thus we will investigate the limits of integrals $\int_0^t \psi_n dB$, where the integrands $\{\psi_n(t); t \in [0, 1]\}_{n=1}^{\infty}$ are supposed to be measurable and adapted to $\{\mathcal{F}(t)\}$ and to satisfy

$$\int_0^1 \psi_n^2 ds < \infty \text{ a.s.}, \qquad n = 1, 2, \cdots.$$

The natural time scale of $\int_0^t \psi_n dB$ is given by

$$\tau_n(t) = \int_0^t \psi_n^2 ds,$$

and we define an inverse to it as $\tau_n^{-1}(s) = \inf\{s \geq 0; \tau_n(s) \geq t\}$.

In this section two functional limit theorems are proved. The main interest of the first one (Theorem 1.1), which requires somewhat weaker conditions, is in the one dimensional limit results which follow from it. Thus an immediate corollary of it is that $\int_0^1 \psi_n dB \to_d \int \Phi(\cdot / s^{\frac{1}{2}}) dF_\tau(s)$, where Φ is the standard normal distribution function and F_τ is the distribution function of the limit τ of $\{\tau_n(1)\}$. Somewhat less obviously it also follows that $\sup_{0 \leq t \leq 1} |\int_0^t \psi_n dB| \to \int G(\cdot / s^{\frac{1}{2}}) dF_\tau(s)$ with $G(x) = P(\sup_{0 \leq t \leq 1} |B(t)| \leq x)$. Further it is possible to get simpler limits by using random normalizations, e.g., on the set $\{\tau > 0\}$, $\int_0^1 \psi_n dB / \tau_n(1)^{\frac{1}{2}} \to_d \Phi$, and similarly for the supremum over $[0, 1]$.

In the results below the functional limits are for random variables in $C(0, 1)$, and vectors of random variables are taken to belong to the product space given the product topology. It should be noted that τ_n is nondecreasing and that it hence converges in probability in $C(0, 1)$ if and only if it converges in probability at each time point separately. To describe the limiting distribution we use another standard Brownian motion $\{W(t); t \geq 0\}$ which is supposed to be independent of everything else.

THEOREM 1.1. *Suppose that*

$$\tau_n(1) = \int_0^1 \psi_n^2 ds \to_p \tau, \qquad n \to \infty, \tag{1.1}$$

for some random variable τ and that

$$\sup_{0 \leq t \leq 1} |\int_0^t \psi_n ds| \to_p 0, \qquad n \to \infty. \tag{1.2}$$

Then

$$\left(\int_0^{\tau_n^{-1}((\cdot)\tau_n(1))} \psi_n dB, \tau_n(1)\right) \to_d (W((\cdot)\tau), \tau), \qquad n \to \infty, \tag{1.3}$$

where the Brownian motion W is independent of τ.

PROOF. To assure that $\int_0^{\tau_n^{-1}(t)} \psi_n dB$ is well defined for all $t \geq 0$, we will during the proof assume that $B(t)$, $\psi_n(t)$, and $\mathcal{F}(t)$ are defined also for $t > 1$, with $\psi_n(t) = 0$, for $1 < t \leq n$, and $\psi_n(t) = 1$, for $t > n$. This assumption involves neither the hypothesis nor the conclusion of the theorem, and can be made without loss of generality. It is known, see, e.g., McKean (1969, page 29) that $\int_0^{\tau_n^{-1}} \psi_n dB$ then is a Brownian motion on $[0, \infty)$.

The main part of the proof is to show that $\int_0^{\tau_n^{-1}}\psi_n dB$ is Rényi-mixing, i.e., that for each probability $\tilde{P}$ which is absolutely continuous with respect to P, the sequence converges in distribution to a Brownian motion W also when P is replaced by $\tilde{P}$ (see, e.g., [8], [9] or [4]). Here convergence in distribution is for random variables in $C(0, \infty)$ given the topology of uniform convergence on compacts (see Whitt (1970)). An alternative way of formulating Rényi-mixing is to require that

$$(1.4) \qquad E\left\{\xi f\left(\int_0^{\tau_n^{-1}}\psi_n dB\right)\right\} \to E(\xi)E(f(W)), \qquad n \to \infty,$$

for continuous bounded functions f: $C(0, \infty) \to R$ if $\xi \geqslant 0$, $E(\xi) < \infty$. It is easy to see that if a random variable ξ' can be approximated arbitrarily well in the mean by variables ξ satisfying (1.4) then (1.4) holds also if ξ is replaced by ξ'. Now it is known, see, e.g., Doob (1953, page 606) that if $E(|\xi'|) < \infty$ then ξ' can be approximated in the mean by $\xi \in \mathcal{F}(T)$ for finite T. Further, for $\xi' \geqslant 0$ a.s. it is no restriction to suppose that the approximating ξ are strictly positive, and after division by $E(\xi)$ it may be assumed that $E(\xi) = 1$.

It thus only remains to be proved that (1.4) holds for $\xi \in \mathcal{F}(T)$ with $\xi > 0$ a.s. and $E(\xi) = 1$. For such ξ the relation

$$\frac{d\tilde{P}}{dP} = \xi$$

defines a probability $\tilde{P}$ which is equivalent to P. From the results of Duncan (1970) it follows that there exists a measurable adapted process $\{\eta(t); t \geqslant 0\}$ with $\eta(t) = 0$ for $t > T$ and with $\int_0^T\eta^2 ds < \infty$ a.s. such that $\xi = \exp\{\int_0^T\eta dB - \frac{1}{2}\int_0^T\eta^2 ds\}$. Thus, by Girsanov's theorem (Girsanov (1960))

$$\tilde{B}(t) = B(t) - \int_0^t\eta ds$$

is a Brownian motion under $\tilde{P}$ and

$$\int_0^t\psi_n dB = \int_0^t\psi_n d\tilde{B} + \int_0^t\psi_n\eta ds.$$

As above, $\int_0^{\tau_n^{-1}}\psi_n d\tilde{B}$ is a Brownian motion under $\tilde{P}$. Thus, since $\eta(t) = 0$ for $t > T$, to prove (1.4) it suffices to show that $\sup_{0\leqslant t\leqslant T}|\int_0^t\psi_n\eta ds| \to_p 0$, as $n \to \infty$, and since $\psi_n(t) = 0$ for $1 < t \leqslant T$ if n is large enough this is the same as proving

$$(1.5) \qquad \sup_{0\leqslant t\leqslant 1}\left|\int_0^t\psi_n\eta ds\right| \to_p 0, \qquad n \to \infty.$$

By (1.2) this clearly holds for η of the form $\eta(t) = \sum_{i=0}^{[kT]-1}\eta_i I\{i/k \leqslant t < (i + 1)/k\}$ and further for arbitrary η with $\int_0^1\eta^2 ds < \infty$ a.s. it is possible to find $\eta^{(k)}$ of this form such that $\int_0^1(\eta - \eta^{(k)})^2 ds \to_p 0$ as $n \to \infty$. Thus, by Cauchy's inequality and (1.1)

$$\sup_{0\leqslant t\leqslant 1}\left|\int_0^t\psi_n\eta ds\right| \leqslant \sup_{0\leqslant t\leqslant 1}\left|\int_0^t\psi_n\eta^{(k)} ds\right| + \left\{\int_0^1\psi_n^2 ds\int_0^1(\eta - \eta^{(k)})^2 ds\right\}^{1/2}$$

$$\to_p 0 + \left\{\tau\int_0^1(\eta - \eta^{(k)})^2 ds\right\}^{1/2}.$$

Letting $k \to \infty$ now proves (1.5) and hence that $\{\int_0^{\tau_n^{-1}}\psi_n dB\}$ is Rényi-mixing.

Since $\tau_n(1) \to_p \tau$ by the assumption (1.1) it follows from Theorem 4.5 of

Billingsley (1968) that

$$\left(\int_0^{\tau_n^{-1}} \psi_n dB, \tau_n(1)\right) \to_d (W, \tau), \qquad n \to \infty,$$

in $C(0, \infty) \times \mathbb{R}$, where W is a Brownian motion independent of τ. Further, by arguments similar to those on page 145 of the cited reference, this implies that

$$\left(\int_0^{\tau_n^{-1}((\cdot)\tau_n(1))} \psi_n dB, \tau_n(1)\right) \to_d (W((\cdot)\tau), \tau) \qquad n \to \infty,$$

in $C(0, 1) \times \mathbb{R}$. □

In a different context, normalizations similar to those above have been considered by Hall (1976). If the first part of the hypothesis of the theorem is strengthened, it is possible to get a functional limit theorem not only for the time-changed integral $\int_0^{\tau_n^{-1}(t\tau_n(1))} \psi_n dB$ but also for the integral $\int_0^t \psi_n dB$ itself.

THEOREM 1.2. *Suppose that*

$$\tau_n(t) = \int_0^t \psi_n^2 ds \to_p \tau(t), \qquad t \in [0, 1], \tag{1.6}$$

as $n \to \infty$, *for some continuous process* $\{\tau(t); t \in [0, 1]\}$ *and that* (1.2) *is satisfied. Then*

$$\int_0^{(\cdot)} \psi_n dB \to_d W \circ \tau, \qquad n \to \infty, \tag{1.7}$$

where the Brownian motion W *is independent of* τ.

PROOF. As proved above $\int_0^{\tau_n^{-1}} \psi_n dB$ is Rényi-mixing so $(\int_0^{\tau_n^{-1}} \psi_n dB, \tau_n) \to_d (W, \tau)$ and hence $\int_0^{(\cdot)} \psi_n dB =_{\text{a.s.}} \int_0^{\tau_n^{-1} \circ \tau_n} \psi_n dB \to_d W \circ \tau$. □

REMARK 1.3. In fact we have proved a little more than the conclusions of Theorems 1.1 and 1.2, namely that the convergence is stable in the sense of Rényi. This can be said in another way, which will be useful in Section 3: if a sequence of processes, say $\{X_n(t); t \in [0, 1]\}$, converges in probability to a process $\{X(t); t \in [0, 1]\}$, then X_n converges *jointly* with the quantities in (1.3) and (1.7). For further information about stability in the sense of Rényi see, e.g., Rényi (1970) or Eagleson (1975). □

Another consequence of stability is that there is no stronger convergence, as can be seen from the following theorem.

THEOREM 1.4. *Suppose that the hypothesis of Theorem* 1.1 *is satisfied. Then on the set* $\{\tau > 0\}$ *the range* $\{\int_0^1 \psi_1 dB, \int_0^1 \psi_2 dB, \cdots\}$ *of* $\int_0^1 \psi_n dB$ *is dense in* $\mathbb{R}$ a.s.

PROOF. We have to prove that $\{\int_0^1 \psi_1 dB, \int_0^1 \psi_2 dB, \cdots\}$ is dense in $\mathbb{R}$ a.s. with respect to $\tilde{P}(\cdot) = P(\cdot | \tau > 0)$. Proceeding as in the proof of Theorem 2.2 of Rootzén (1976) we only have to show that

$$\limsup_{n \to \infty} \tilde{P}\left(\int_0^1 \psi_n dB \in (x - \varepsilon, x + \varepsilon] | A\right) > 0$$

if $\tilde{P}(A) > 0$ and $\varepsilon > 0$, $x \in \mathbb{R}$. However, above it was proved that $(\int_0^{\tau_n^{-1}} \psi_n dB, \tau_n)$ $\to_d (W, \tau)$ also under $\tilde{P}(\cdot | A) = P(\cdot | A \cap \{\tau > 0\})$ and hence, putting $\tilde{F}(x) = P(\tau$

$\leqslant x | A \cap \{\tau > 0\}),$

$$\tilde{P}\big(\int_0^1 \psi_n dB \in (x - \varepsilon, x + \varepsilon] | A\big) \to \int \big\{\Phi\big((x + \varepsilon)/s^{\frac{1}{2}}\big) - \Phi\big((x - \varepsilon)/s^{\frac{1}{2}}\big)\big\} d\bar{F}(s) > 0.$$

□

Results similar to those of Theorem 1.4 also hold for $\sup_{0 \leqslant t \leqslant 1} |\int_0^t \psi_n dB|$, for $\int_0^1 \psi_n dB / \tau_n(1)^{\frac{1}{2}}$, etc., and it can be seen that they still hold if n is restricted to some (infinite) subset of $\{1, 2, \cdots\}$. Moreover, if the hypothesis of Theorem 1.2 is satisfied, then analogous results hold for the entire process $\int_0^t \psi_n dB$; $t \in [0, 1]$. For instance, on the set $\{\omega;\ \tau(t) > 0 \text{ for } t > 0\}$, the $C(0, 1)$-closure of $\{\int_0^{(\cdot)} \psi_1 dB, \int_0^{(\cdot)} \psi_2 dB, \cdots\}$ then is a.s. the set of $C(0, 1)$ functions which are zero at zero.

We end this section with a lemma which will be useful when checking (1.2) and (1.6) in later sections. In the lemma $I(\cdot)$, as is customary, denotes an indicator function, i.e., $I(A)$ is one on A and zero on A^c.

Lemma 1.5. *Suppose that the stochastic process* $\{a(t);\ t \in [0, 1]\}$ *is* a.s. *Rieman integrable over* $[0, 1]$, *let* $\psi_n(t) = n^{k/2} \sum_{i=0}^{n-1} a(i/n)(B(t) - B(i/n))^k I(i/n \leqslant t < (i + 1)/n)$ *and write* $E_k = E\int_0^1 B(s)^k ds$. *Then*

$$\sup_{0 \leqslant t \leqslant 1} \Big|\int_0^t \psi_n ds - E_k \int_0^t a\, ds\Big| \to_p 0, \qquad n \to \infty, \tag{1.8}$$

for $k > 0$.

Proof. Put $b_n(t) = n^{k/2} \sum_{i=0}^{n-1} (B(t) - B(i/n))^k I(i/n \leqslant t < (i + 1)/n)$. Clearly $\int_{i/n}^{(i+1)/n} (B(s) - B(i/n))^k ds$ has the same distribution as $n^{-1-k/2} \int_0^1 B(s)^k ds$ and hence $E\{\int_0^{i/n} b_n ds - E_k i/n\} = 0$, $i = 1, \cdots, n$, and $E\{\int_0^1 b_n ds - E_k\}^2 \to 0$ as $n \to \infty$. Thus, by Kolmogorov's inequality,

$$\max_{1 \leqslant i \leqslant n} \Big|\int_0^{i/n} b_n ds - E_k i/n\Big| \to_p 0, \qquad n \to \infty.$$

Further, using that $E\int_0^1 b_n^2 ds$ is bounded in n, Cauchy's inequality gives that

$$\sup_{i/n \leqslant t < (i+1)/n} \Big|\int_{i/n}^t b_n ds\Big| \leqslant \big\{n^{-1} \textstyle\int_0^1 b_n^2 ds\big\}^{1/2} \to_p 0, \qquad n \to \infty,$$

and hence

$$\sup_{0 \leqslant t \leqslant 1} \Big|\int_0^t b_n ds - E_k t\Big| \to_p 0, \qquad n \to \infty. \tag{1.9}$$

Writing $a_n(t) = \sum_{i=0}^{n-1} a(i/n) I(i/n \leqslant t < (i + 1)/n)$ we have $\psi_n(t) = a_n(t) b_n(t)$, and if $a(t)$ is a stepfunction it follows at once that (1.8) holds. The general case is then proved by approximating a by stepfunctions and then using Cauchy's inequality and that $E\int_0^1 b_n^2 ds$ is bounded in n (cf. the proof of (1.5) above). □

It may be noted that the lemma easily can be extended to deal with other partitions of [0, 1] than 0, $1/n$, $2/n$, $\cdots$, 1 and that (1.2) can be checked for more general ψ_n's by requiring that ψ_n is adapted to $\{\mathcal{F}(t)\}$.

2. A special case. The results from the previous section will now be illustrated by considering approximations of the integral $\int_0^t f(B(s), s)dB(s)$. To avoid unnecessary complications we study the most basic case; when $f(B(t), t)$ is approximated by $\Sigma_{i=0}^{n-1} f(B(i/n), i/n)I(i/n \leqslant t < (i+1)/n)$. Thus, writing Δ_n^i for $B((i+1)/n) - B(i/n)$ if $i < [nt]$, and for $B(t) - B(i/n)$ if $i = [nt]$, we are interested in the deviation of $\int_0^t f(B(s), s)dB(s)$ from $\Sigma_{i=0}^{[nt]} f(B(i/n), i/n)\Delta_n^i$. It is possible to treat other approximation schemes in a quite similar way. In the next section two such schemes are studied.

THEOREM 2.1. *Suppose that* $f(x, t) - f(x, s) = o(|t - s|^{\frac{1}{2}})$ *uniformly on compacts, and that the partial* $\frac{\partial}{\partial x} f(x, t) = f_1(x, t)$ *is continuous. Then*

$$\left\{ n^{\frac{1}{2}}\left(\int_0^t f(B(s), s)dB(s) - \Sigma_{i=0}^{[nt]} f(B(i/n), i/n)\Delta_n^i\right);\ t \in [0, 1] \right\} \to_d W \circ \tau, \tag{2.1}$$

where the Brownian motion W is independent of τ and where $\tau(t) = \frac{1}{2}\int_0^t f_1(B(s), s)^2 ds$.

PROOF. Put $\phi(t) = f(B(t), t)$, $\phi_n(t) = \Sigma_{i=0}^{n-1} f(B(i/n), i/n)I(i/n \leqslant t < (i+1)/n)$ and let $\psi_n(t) = n^{\frac{1}{2}}\Sigma_{i=0}^{n-1} f_1(B(i/n), i/n)(B(t) - B(i/n))I(i/n \leqslant t < (i+1)/n)$. The first step of the proof is to approximate $n^{\frac{1}{2}}\int_0^t(\phi - \phi_n)dB$ by $\int_0^t \psi_n dB$. Now, writing $I_n(i) = [i/n, (i+1)/n)$,

$$\begin{aligned} n^{\frac{1}{2}}\int_0^t(\phi - \phi_n)dB &= n^{\frac{1}{2}}\Sigma_{i=0}^{n-1} \int_{I_n(i)\cap[0,t]}(f(B(s), s) - f(B(s), i/n))dB \\ &\quad + \int_0^t \psi_n dB + n^{\frac{1}{2}}\Sigma_{i=0}^{n-1} \int_{I_n(i)\cap[0,t]} o(|B(s) - B(i/n)|)dB \\ &= R_n^1(t) + \int_0^t \psi_n dB + R_n^2(t), \quad \text{say}. \end{aligned} \tag{2.2}$$

To prove that R_n^1 and R_n^2 converge uniformly to zero in probability we will use the fact that if, for a sequence $\{e_n(t);\ t \in [0, 1]\}$ of integrands, $\int_0^1 e_n^2 ds \to_p 0$ then $\sup_{0 \leqslant t \leqslant 1}|\int_0^t e_n dB| \to_p 0$. Since almost every sample path of $\{B(t), t \in [0, 1]\}$ is bounded we have from the first condition of the theorem that

$$\begin{aligned} n\, \Sigma_{i=0}^{n-1} \int_{I_n(i)}(f(B(s), s) - f(B(s), i/n))^2 ds & \\ = o(1)n\, \Sigma_{i=0}^{n-1} \int_{I_n(i)}(s - i/n)ds \to_{\text{a.s.}} 0, & \end{aligned}$$

as $n \to \infty$. Similarly,

$$\begin{aligned} n\, \Sigma_{i=0}^{n-1} \int_{I_n(i)} o(|B(s) - B(i/n)|)^2 ds = o(1)n\, \Sigma_{i=0}^{n-1} \int_{I_n(i)}(B(s) - B(i/n))^2 ds \to \\ \to_p 0, \qquad n \to \infty, \end{aligned}$$

since $o(1) \to_{\text{a.s.}} 0$ and since $\int_{I_n(i)}(B(s) - B(i/n))^2 ds$ has the same distribution as $n^{-2}\int_0^1 B(s)^2 ds$ and $\{n\Sigma_{i=0}^{n-1}\int_{I_n(i)}(B(s) - B(i/n))^2 ds\}$ hence is tight. Thus we have proved

$$\sup_{0 \leqslant t \leqslant 1}|R_n^i(t)| \to_p 0 \tag{2.3}$$

as $n \to \infty$ for $i = 1, 2$.

Next, by Lemma 1.5 with $a(t)$ replaced by $f_1(B(t), t)$ and $f_1(B(t), t)^2$, respectively, which are continuous and hence Rieman integrable,

$$\sup_{0 \leqslant t \leqslant 1} |\int_0^t \psi_n ds| \to_p 0, \qquad n \to \infty,$$

and

$$\int_0^t \psi_n^2 ds \to_p \tfrac{1}{2} \int_0^t f_1(B(s), s)^2 ds, \; t \in [0, 1], \qquad n \to \infty,$$

since clearly $E_1 = 0$ and $E_2 = \frac{1}{2}$. Thus $\{\psi_n\}$ satisfies the hypothesis of Theorem 1.2 and hence

$$\int_0^{(\cdot)} \psi_n dB \to_d W \circ \tau, \qquad n \to \infty,$$

and by (2.2) and (2.3) this proves the theorem. □

3. Stratonovich integrals and higher order approximations. The methods developed above can also be used to study convergence to the Stratonovich (1964) integral and convergence of the Rieman-Stieltjes approximations considered by Wong and Zakai (1965). Furthermore, from the preceding section can be seen that higher order approximation schemes will converge much quicker to the Itô integral, and that with a suitable normalization it is possible to get nontrivial limit distributions also for such approximations. The necessary computations follow the same lines as before, and hence they will only be sketched. Moreover, for simplicity only integrands $f(B(s))$ with f four times continuously differentiable will be treated.

In Yor (1976) is used the following convenient formulation, which includes both the Stratonovich integral and the Rieman-Stieltjes approximations. Let μ be a probability measure on $([0, 1], \mathfrak{B}([0, 1]))$ with $\mu_k = \int_0^1 x^k d\mu(x)$, $k = 1, 2, \cdots$ and, using the notation Δ_n^i from the previous page, put

$$I_n(t) = \Sigma_{i=0}^{[nt]} \int_0^1 f(B(i/n) + s\Delta_n^i) d\mu(s) \Delta_n^i.$$

In the cited paper it is proved that

$$I_n(t) \to_p \int_0^t f(B(s)) dB(s) + \mu_1 \int_0^t f'(B(s)) ds$$

as $n \to \infty$. Here the Itô integral corresponds to $\mu = \delta_0$, the Stratonovich integral to $\mu = \delta_{\frac{1}{2}}$, and the Rieman-Stieltjes approximation obtained by making a polygonal approximation of $B(s)$ between the points $0, 1/n, 2/n, \cdots n/n$ corresponds to μ being Lebesgue measure on [0, 1].

Let

$$d_n(t) = \int_0^t f(B) dB + \mu_1 \int_0^t f'(B) ds - I_n(t). \tag{3.1}$$

It turns out that the behaviour of d_n is quite different according to whether $\mu_1 \neq \frac{1}{2}$ or $\mu_1 = \frac{1}{2}$. In the former case the proper normalizing constants are $n^{\frac{1}{2}}$ while in the latter case one should use n or higher powers of n. For the Stratonovich integral and for the Rieman-Stieltjes approximations $\mu_1 = \frac{1}{2}$, but of course in the latter case $I_n(t)$ does not depend on n and hence the approximation error is zero. However, for integrands $f(B(s), s)$ there is a nonzero approximation error (typically of the order

$1/n$) also in this case. For smooth integrands $f(B(s))$ the order of the approximation error in Yor's scheme is determined by the first k for which $\mu_k \neq 1/(1+k)$.

We first show that

$$n^{\frac{1}{2}} d_n \rightarrow_d W \circ \tau, \tag{3.2}$$

as $n \rightarrow \infty$, where d_n is given by (3.1) and where the Brownian motion W is independent of τ and $\tau(t) = \frac{1}{2}(1 - 2\mu_1)^2 \int_0^t f'(B(s))^2 ds$. If $\mu_1 = \frac{1}{2}$, this distribution is degenerate, and later we will consider the limiting distribution of $n\, d_n$ in this case.

The calculations which follow are longish, so it is helpful to abbreviate the notation. In most places we delete explicit dependence on n, and hence, if $i < [nt]$, we write Δ_i for $B((i+1)/n) - B(i/n)$ and I_i for the interval $[i/n, (i+1)/n)$. If $i = [nt]$ then $\Delta_i = B(t) - B(i/n)$ and $I_i = [i/n, t)$. Further $|I_i|$ denotes the length of the ith interval, $\Delta_i(s) = B(s) - B(i/n)$ and B_i is short for $B(i/n)$, f_i for $f(B(i/n))$, f_i' for $f'(B(i/n))$, etc. A sequence of stochastic processes, say $\{e_n(t);\, t \in [0, 1]\}_{n=1}^{\infty}$, is $o_p(f(n))$ if $\sup_{0 \leq t \leq 1} |X_n(t)|/f(n) \rightarrow_p 0$ as $n \rightarrow \infty$.

With this notation, expanding the last two terms of $d_n(t)$ around $B_i (= B(i/n))$ gives

$$\begin{aligned} d_n(t) = {} & \Sigma_{i=0}^{[nt]} \{ \textstyle\int_{I_i} f(B) dB + \mu_1 f_i' |I_i| - f_i \Delta_i - \mu_1 f_i' \Delta_i^2 \} \\ & + \Sigma_{i=0}^{[nt]} \{ \mu_1 \textstyle\int_{I_i} (f'(B(s)) - f_i') ds - \int_0^1 (f(B_i + s\Delta_i) - f_i - f_i' s \Delta_i) d\mu(s) \Delta_i \}. \end{aligned}$$

The methods of Section 2 can be used to show that the latter sum is $o_p(n^{-\frac{1}{2}})$. Thus, by expanding around B_i also in the first integral above, and by using the fact that $\Delta_i^2 - |I_i| = 2\int_{I_i} \Delta_i(s) dB(s)$ we obtain

$$d_n(t) = (1 - 2\mu_1) \Sigma_{i=0}^{[nt]} \textstyle\int_{I_i} f_i' \Delta_i(s) dB(s) + o_p(n^{-\frac{1}{2}}). \tag{3.3}$$

From Lemma 1.5 and Theorem 1.2 it follows that

$$\left\{ n^{\frac{1}{2}} (1 - 2\mu_1) \Sigma_{i=0}^{[nt]} \textstyle\int_{I_i} f_i' \Delta_i(s) dB(s);\, t \in [0, 1] \right\} \rightarrow_d W \circ \tau,$$

which together with (3.3) proves (3.2).

Next we assume $\mu_1 = \frac{1}{2}$ and show that in this case

$$n\, d_n \rightarrow_d W \circ \tau + \tfrac{1}{2}(1 - 3\mu_2) \textstyle\int_0^t f''(B(s)) dB(s) + \tfrac{1}{8}(1 - 4\mu_3) \int_0^t f'''(B(s)) ds, \tag{3.4}$$

where the Brownian motion W is independent of B and where $\tau(t) = \frac{1}{6}(1 - 3\mu_2)^2 \int_0^t f''(B(s))^2 ds$. This limiting distribution is degenerate if $\mu_2 = \frac{1}{3}$ and $\mu_3 = \frac{1}{4}$ (still assuming $\mu_1 = \frac{1}{2}$), which, as it should, holds for the Rieman-Stieltjes approximation, i.e., for μ being Lebesgue measure.

When $\mu_1 = \frac{1}{2}$ the low order terms in the expansion of $d_n(t)$ cancel, and the next higher order terms have to be retained. There is an additional minor complication. A straightforward second order expansion of $\Sigma_{i=1}^{[nt]} \int_{I_i} f'(B(s)) ds$ gives the "remainder" $\Sigma_{i=1}^{[nt]} \int_{I_i} [f'(B(s)) - f_i' - f_i'' \Delta_i(s)] ds = \Sigma_{i=1}^{[nt]} \int_{I_i} f_i''' \Delta_i(s)^2/2\, ds + o_p(n^{-1})$, but here $E(\Delta_i(s)^2) = s - i/n\ (> 0)$ and it follows that the sum is not $o_p(n^{-1})$ as it should be. However, this can be remedied by subtracting $\Sigma_{i=1}^{[nt]} f_i''' |I_i|^2/4 =$

$\Sigma_{i=1}^{[nt]}\int_{I_i}f_i'''(s - i/n)/2ds$, since the remainder then becomes $\Sigma_{i=1}^{[nt]}\int_{I_i}f_i'''[\Delta_i(s)^2 - (s - i/n)]/2ds + o_p(n^{-1})$, which is $o_p(n^{-1})$. Together with a similar adjustment by $\Sigma_{i=0}^{[nt]}f_i'''|I_i|^2/2$ in the expansion of $\Sigma_{i=0}^{[nt]}\int_{I_i}f(B_i + s\Delta_i)d\mu(s)\Delta_i$, it ensues that

$$d_n(t) = \Sigma_{i=0}^{[nt]}\Big\{\int_{I_i}f_i''\big(\Delta_i(s)^2/2\big)dB(s) + \mu_1\int_{I_i}f_i''\Delta_i(s)ds + \mu_1 f_i'''|I_i|^2/4 - \mu_2 f_i''\Delta_i^3/2 - \mu_3 f_i'''|I_i|^2/2\Big\} + o_p(n^{-1}).$$

Further $\Delta_i^3 = 3\int_{I_i}\Delta_i(s)^2dB(s) + 3\int_{I_i}\Delta_i(s)ds$, and by inserting $\mu_1 = \frac{1}{2}$ and rearranging terms we obtain

$$d_n(t) = \Sigma_{i=0}^{[nt]}\Big\{\big(\tfrac{1}{2} - 3\mu_2/2\big)\int_{I_i}f_i''\Delta_i(s)^2dB(s) + \big(\tfrac{1}{2} - 3\mu_2/2\big)\int_{I_i}f_i''\Delta_i(s)ds + \big(\tfrac{1}{8} - \mu_3/2\big)f_i'''|I_i|^2\Big\} + o_p(n^{-1}).$$

Partial integration gives that $\int_{I_i}\Delta_i(s)ds = \Delta_i/n - \int_{I_i}(s - i/n)dB(s)$, and hence

$$d_n(t) = \Sigma_{i=0}^{[nt]}\Big\{\big(\tfrac{1}{2} - 3\mu_2/2\big)\int_{I_i}f_i''\big[\Delta_i(s)^2 - (s - i/n)\big]dB(s) + \big(\tfrac{1}{2} - 3\mu_2/2\big)f_i''\Delta_i/n + \big(\tfrac{1}{8} - \mu_3/2\big)f_i'''|I_i|^2\Big\} + 0_p(n^{-1}). \tag{3.5}$$

Lemma 1.5 and Theorem 1.2 can now be applied to prove that

$$\Big\{n\Sigma_{i=0}^{[nt]}\big(\tfrac{1}{2} - 3\mu_2/2\big)\int_{I_i}f_i''\big[\Delta_i(s)^2 - (s - i/n)\big]dB(s);\ t \in [0, 1]\Big\} \to_d W \circ \tau. \tag{3.6}$$

Since f''' (and hence f'') are assumed to be continuous,

$$n\Sigma_{i=0}^{[nt]}\big(\tfrac{1}{2} - 3\mu_2/2\big)f_i''\Delta_i/n \to \tfrac{1}{2}(1 - 3\mu_2)\int_0^t f''(B(s))dB(s) \tag{3.7}$$

and

$$n\Sigma_{i=0}^{[nt]}\big(\tfrac{1}{8} - \mu_3/2\big)f_i'''|I_i|^2 \to \tfrac{1}{8}(1 - 4\mu_3)\int_0^t f'''(B(s))ds \tag{3.8}$$

uniformly in $t \in [0, 1]$, in probability. Now, by Remark 1.3, the expressions (3.5)-(3.8) prove that (3.4) holds.

If one knows the derivative of f, then a straightforward attempt to improve the approximation of $\int_0^t f(B)dB$ by $\Sigma_{i=0}^{[nt]}f(B(i/n))\Delta_i = \Sigma_{i=0}^{[nt]}\int_{I_i}f(B(i/n))dB$ is to replace the latter quantity by $\Sigma_{i=0}^{[nt]}\int_{I_i}[f(B(i/n)) + f'(B(i/n))\Delta_i(s)]dB(s) = \Sigma_{i=0}^{[nt]}\{f(B(i/n))\Delta_i + \frac{1}{2}f'(B(i/n))(\Delta_i^2 - |I_i|)\}$. The approximation error then decreases from the order $n^{-\frac{1}{2}}$ to the order n^{-1}, and in fact we have that

$$\Big\{n\Big(\int_0^t f(B)dB - \Sigma_{i=0}^{[nt]}\big[f(B(i/n))\Delta_i + \tfrac{1}{2}f'(B(i/n))\big(\Delta_i^2 - |I_i|\big)\big]\Big);\ t \in [0, 1]\Big\} \to_d W \circ \tau + \tfrac{1}{4}\int_0^t f''(B(s))dB(s), \tag{3.9}$$

where W is independent of B and where $\tau(t) = \frac{3}{16}\int_0^t f''(B(s))^2ds$. The proof of (3.9) follows the same lines as above, after writing

$$\begin{aligned}\int_0^t f(B)dB &- \Sigma_{i=0}^{[nt]}\int_{I_i}(f_i + f_i'\Delta_i(s))dB(s)\\ &= \Sigma_{i=0}^{[nt]}\int_{I_i}f_i''\Delta_i(s)^2/2dB + o_p(n^{-1})\\ &= \Sigma_{i=0}^{[nt]}\tfrac{1}{2}\int_{I_i}f_i''\big(\Delta_i(s)^2 - |I_i|/4\big)dB + \Sigma_{i=0}^{[nt]}(|I_i|/2)f_i''\Delta_i + o_p(n^{-1}).\end{aligned}$$

It is now of course quite immediate to find limiting distributions also for approximations of higher order than two. An additional term, corresponding to $4^{-1}\int_0^t f''(B(s))dB(s)$ in (3.9) only arises in even order approximations. Further, to consider integrands of the form $f(B(t), t)$ and more elaborate numerical schemes does not seem to introduce any new complications.

Acknowledgment. I want to thank the referee for comments which led to considerable improvement of the presentation of this paper.

REFERENCES

[1] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[3] Duncan, T. E. (1970). On the absolute continuity of measures. *Ann. Math. Statist.* **41** 30–38.

[4] Eagleson, G. K. (1975). The martingale central limit theorem and random norming. Preprint, University of Cambridge.

[5] Girsanov, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theor. Probability Appl.* **5** 285–301.

[6] Hall, P. (1976). Martingale invariance principles. Preprint, Mathematical Institute, University of Oxford.

[7] McKean, H. P. (1969). *Stochastic Integrals*. Academic Press, New York.

[8] Rényi, A. (1970). *Foundations of Probability*. Holden-Day, San Francisco.

[9] Rootzén, H. (1976). Fluctuations of sequences which converge in distribution. *Ann. Probability*. 456–463.

[10] Stratonovich, R. L. (1966). A new representation for stochastic integrals. *Vestnik Moskov Univ. Ser.* I., **1** 3–12 (1964). Reprinted in *SIAM J. of Control* **4** 362–371.

[11] Whitt, W. (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.* **41** 939–944.

[12] Wong, E. and Zakai, M. (1965). On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.* **36** 1560–1564.

[13] Yor, M. (1976). Sur quelques approximations d'integrales stochastiques. Seminar de Probabilités XI, Springer lecture notes in mathematics No 581.

University of Copenhagen
Institute of Mathematical Statistics
5, Universitetsparken
DK-2100 Copenhagen Ø, Denmark