CONTINUOUS MARKOV PROCESSES AND STOCHASTIC EQUATIONS

By **Gisirō Maruyama** (Tokio)

1. *Introduction.* Let $y(t)$ be a Markov process depending on a time parameter t, $0 \leqq t \leqq 1$, with the transition probabilities $F(s, x; t, y) \equiv \Pr. \{y(t) \leqq y \mid y(s) = x\}$. Kolmogorov's celebrated paper [14] laid down a foundation for studying the process and Feller [6] [7] amplified Kolmogorov's method. These theories are based on functional equations satisfied by $F(s, x; t, y)$, i.e. partial differential equations or integro-differential equations etc. In this paper we restrict ourseles to those types of processes which satisfy the conditions: when $t \to s$, there exist the limits

$$(C_1) \qquad \frac{1}{t-s} \int_{|y-x|>\delta} dF(s, x; t, y) \to 0,$$

$$(C_2) \qquad \frac{1}{t-s} \int_{|y-x|\leq\delta} (y-x)\, dF(s, x; t, y) \to a(s, x),$$

$$(C_3) \qquad \frac{1}{t-s} \int_{|y-x|\leq\delta} (y-x)^2\, dF(s, x; t, y) \to b^2(s, x).$$

These conditions have been introduced by Kolmogorov and Feller as ones giving the stochastic process continuity, in some sense or other, and distinguishing it from others which obey discontinuous changes. The corresponding functional equation then becomes a partial differential equation of parabolic type. Under certain restrictions imposed on $a(s, x)$, $b(t, x)$, Feller constructed its fundamental solution. The precise meaning of continuity resulting from (C_1)-(C_3) was not clear until Fortet [8] first proved, using the fundamental solution, that the process determined by the parabolic equation has the path functions continuous with probability 1, and indeed they have the same modulus of continuity as the Brownian motion. The continuity was observed in an independent and different approach by K. Itō which will be stated in the following.

Another important approach to the same problem we owe S. Bernstein [1] in which he introduced a stochastic difference equation and showed that the random variable determined by this equation has the fundamental solution of a Fokker-Planck equation as its limiting distribution. In this direction a remarkable improvement has been made recently by Itō [9] [10], including the mixed type of processes [11]. Itō's method is based on his theory of the stochastic integral with respect to fundamental stochastic processes, e.g. the Brownian motion or the Poisson process. In our continuous case, Itō's main theorem reads: If $a(t, x)$, $b(t, x)$ are continuous in t, x and satisfy the Lipschitz condition

$$|a(t, x) - a(t, x')| + |b(t, x) - b(t, x')| < c|x - x'|, \tag{1.1}$$

then the stochastic integral equation

$$y(t) = y_0 + \int_0^t a[\tau, y(\tau)]\,d\tau + \int_0^t b[\tau, y(\tau)]\,dx(\tau), \tag{1.2}$$

$x(t)$ the Brownian motion (Wiener process) with $x(0) = 0$, $E(\Delta x(t))^2 = \Delta t$, has the unique solution $y(t)$, continuous with probability 1, and it is a Markov process satisfying (C_1)-(C_3).

To study relations between these different approaches, of which one depends on functional equations, and the other on measure in function spaces is an important problem. This paper deals with this problem and presents details of the results announced in the preliminary report [15].

2. *Stochastic difference equation.* To solve (1.2) Itō used a method of successive approximation. If we write it in the form of difference equations we obtain the one introduced by Bernstein. This method provides a finite number of operations to approximate the solution of (1, 2) and proves to be useful for various objects.

THEOREM 1. *Let $x(t)$ be the Brownian motion, $x(0) = 0$, $E(\Delta x(t))^2 = \Delta t$, of which Gaussian increments $\Delta x(t)$ over disjoint intervals are independent, and let $a(t, x)$, $b(t, x)$ be continuous functions satisfying the Lipschitz condition* (1.1). *Consider a division of* $(0, 1)$, $\Delta = \Delta(t_0, t_1, \ldots, t_n)$, $0 = t_0 < t_1 < \cdots < t_n = 1$, *and variables $y_1, y_2, \ldots, y_n$ defined by*

$$\left\{\begin{array}{l} y_1 = y_0 + a(t_0, y_0)\Delta t_0 + b(t_0, y_0)\Delta x_0, \\ y_2 = y_1 + a(t_1, y_1)\Delta t_1 + b(t_1, y_1)\Delta x_1, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ y_n = y_{n-1} + a(t_{n-1}, y_{n-1})\Delta t_{n-1} + b(t_{n-1}, y_{n-1})\Delta x_{n-1} \end{array}\right. \tag{2.1}$$

and also $y_\Delta(t)$ *defined by*

$$y_\Delta(t) = y_\mu + a(t_\mu, y_\mu)(t - t_\mu) + b(t_\mu, y_\mu)[x(t) - x(t_\mu)], \tag{2.2}$$

where

$$x_\nu = x(t_\nu), \quad \Delta x_{\nu-1} = x_\nu - x_{\nu-1}, \quad \Delta t_{\nu-1} = t_\nu - t_{\nu-1} \quad t_\mu \leqq t < t_{\mu+1}.$$

Then, for every t, *in the* L^2*-sense, there exists*

$$y(t) = \underset{\rho(\Delta)\to 0}{\mathrm{l.i.m.}}\ y_\Delta(t), \quad \rho(\Delta) = \underset{1\leq\nu\leq n}{\mathrm{Max}}\ \Delta t_{\nu-1}$$

and $y(t)$ *becomes the unique solution of* (1.2).

Throughout this paper we denote by $c(x)$ a function bounded in every finite x-interval, and by c, c', c_1, c_2, ... constants, but not necessarily the same for every case. The initial value random variable y_0 is supposed, without loss of generality, to be a constant, general statements being obtained by taking average with given y_0-distribution.

LEMMA 1. *If* $a(t, x)$, $b(t, x)$ *are continuous functions of* t, x (*but not necessarily satisfy the condition* (1.1)) *and satisfy*

$$|a(t, x)| + |b(t, x)| \leqq c'(1 + x^2)^{1/2} \tag{2.3}$$

then

$$\left.\begin{matrix} E(y_m^{2p}) \\ E(y_\Delta^{2p}(t)) \end{matrix}\right\} \leqq c''(1 + y_0^2)^p, \qquad p = 1, 2, \ldots, \tag{2.4}$$

c'' *being independent of* m, t, y_0, $1 \leqq m \leqq n$.

Proof. (2.1) gives

$$E(y_n^2) = E(y_{n-1}^2) + E\{a^2(t_{n-1}, y_{n-1})\}(\Delta t_{n-1})^2$$
$$+ E\{b^2(t_{n-1}, y_{n-1})\}\Delta t_{n-1} + 2E\{y_{n-1}a(t_{n-1}, y_{n-1})\}\Delta t_{\nu-1}.$$

If we write $\mu_n = E(y_n^2)$ and use (2.3) we get from the above

$$\mu_\nu \leqq \mu_{\nu-1}(1 + c_1 \Delta t_{\nu-1}) + c_1 \Delta t_{\nu-1}, \qquad \nu = 1, 2, \ldots, n.$$

Successive substitution now gives

$$\mu_m \leqq \mu_0^2 \prod_{\nu=0}^{m-1}(1 + c_1 \Delta t_\nu) + \prod_{\nu=0}^{m-1}(1 + c_1 \Delta t_\nu)\sum_{\nu=0}^{m-1}\Delta t_\nu \leqq c_2(1 + y_0^2),$$

since $\mu_0 = y_0$. The same is true with $E(y_\Delta^2(t))$. These prove the lemme for $p = 1$. The general case is obtained inductively by taking average on both sides of the p^{th} power of (2.1).

Proof of Theorem 1. Consider any two divisions of $(0, 1)$, $\Delta = \Delta(t_0, t_1, \ldots, t_m)$, $\Delta' = \Delta(t'_0, t'_1, \ldots, t'_n)$ and denote the corresponding expressions for (2.2) respectively by $y_\Delta(t)$ and $y_{\Delta'}(t)$. Union of the two divisions yields a new division $\Delta(\tau_0, \tau_1, \ldots)$ and by suitably chosen $u_0, u_1, u_2, \ldots, u'_0, u'_1, \ldots, y_\Delta(t), y_{\Delta'}(t)$ can be written as

$$\text{(2.5)}\quad \begin{cases} y_\Delta(t) = y_0 + \sum_{(t)} a[u_{\nu-1}, y_\Delta(u_{\nu-1})]\Delta\tau_{\nu-1} + \sum_{(t)} b[u_{\nu-1}, y_\Delta(u_{\nu-1})]\Delta x_{\nu-1} \\ y_{\Delta'}(t) = y_0 + \sum_{(t)} a[u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1})]\Delta\tau_{\nu-1} + \sum_{(t)} b[u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1})]\Delta x_{\nu-1}, \end{cases}$$

where $\Delta x_{\nu-1} = x(\tau_\nu) - x(\tau_{\nu-1})$ and $\sum_{(t)}$ means that summations are taken over $u_\nu \leqslant t$ and $u'_\nu \leqslant t$, and in the last term differences $\Delta\tau_{\nu-1}$, $\Delta x_{\nu-1}$ stand for $t - t_\mu$, $x(t) - x(t_\mu)$ respectively. We have by (2.5)

$$\text{(2.6)}\quad \begin{cases} E[y_\Delta(t) - y_{\Delta'}(t)]^2 \leqslant 2E\left\{\sum_{(t)}[a(u_{\nu-1}, y_\Delta(u_{\nu-1})) - a(u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1}))]\Delta\tau_{\nu-1}\right\}^2 \\ \qquad + 2E\left\{\sum_{(t)}[b(u_{\nu-1}, y_\Delta(u_{\nu-1})) - b(u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1}))]\Delta x_{\nu-1}\right\}^2 \\ \qquad \leqslant 2\sum_{(t)} E\{a(u_{\nu-1}, y_\Delta(u_{\nu-1})) - a(u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1}))\}^2 \Delta\tau_{\nu-1} \\ \qquad + 2\sum_{(t)} E\{b(u_{\nu-1}, y_\Delta(u_{\nu-1})) - b(u'_{\nu-1}, y_{\Delta'}(u'_{\nu-1}))\}^2 \Delta\tau_{\nu-1}. \end{cases}$$

Now, in general, we have by (1.1)

$$\text{(2.7)}\quad \begin{cases} [a(s, y_\Delta(s)) - a(t, y_{\Delta'}(t))]^2 \leqslant c\{[y_\Delta(t) - y_\Delta(s)]^2 \\ + [y_\Delta(t) - y_{\Delta'}(t)]^2 + [a(s, y_\Delta(t)) - a(t, y_\Delta(t))]^2\} \end{cases}$$

and similarly for b.

To estimate the right-hand member of (2.6) we put

$$F(t, y) = \Pr.\ \{y_\Delta(t) \leqslant y\}$$

and write

$$E\{a[s, y_\Delta(t)] - a[t, y_\Delta(t)]\}^2 = \left(\int_{|y| \geq K} + \int_{|y| < K}\right)[a(s, y) - a(t, y)]^2 dF(t, y) = J_1 + J_2,$$

say. If a, b satisfy (1.1), (2.3) holds and hence by Lemma 1, when an arbitrary $\varepsilon > 0$ is given, we can find $K > 0$ such that

$$\begin{aligned} |J_1| &\leqslant c\int_{|y| > K}(1 + y^2)\, dF(t, y) \\ &\leqslant \frac{c}{1 + K^2}\int_{-\infty}^{\infty}(1 + y^2)^2 dF \leqslant \frac{c(y_0)}{1 + K^2} \leqslant \frac{\varepsilon}{2}. \end{aligned}$$

Also by continuity of a, we can choose $\delta > 0$ such that $|t-s| < \delta$ implies

$$|J_2| < \varepsilon/2.$$

Hence we have

$$E\{a[s, y_\Delta(t)] - a[t, y_\Delta(t)]\}^2 < \varepsilon \cdot c(y_0), \tag{2.8}$$

if $\rho(\Delta) < \delta$, and similary for b. (2.2) also gives us

$$E\{y_\Delta(s) - y_\Delta(t)\}^2 \leqslant c(1+y_0^2)|t-s|. \tag{2.9}$$

Substituting (2.7) into (2.6) combined with (2.8), (2.9) we get

$$\left\{\begin{aligned} \Delta(t) &= E\{[y_\Delta(t) - y_{\Delta'}(t)]^2\} \leqslant c(y_0)\Big\{\sum\nolimits_{(t)}|u_{\nu-1} - u'_{\nu-1}|\Delta\tau_{\nu-1} \\ &\quad + \sum\nolimits_{(t)} E[y_\Delta(u'_{\nu-1}) - y_{\Delta'}(u'_{\nu-1})]^2\Delta\tau_{\nu-1} + \varepsilon\sum\nolimits_{(t)}\Delta\tau_{\nu-1}\Big\} \\ &\leqslant c_1(y_0)\varepsilon + c(y_0)\sum\nolimits_{(t)} E[y_\Delta(u'_{\nu-1}) - y_{\Delta'}(u'_{\nu-1})]^2\Delta\tau_{\nu-1}, \\ &\qquad \rho(\Delta),\quad \rho(\Delta') < \delta. \end{aligned}\right. \tag{2.10}$$

As will easily be shown $\Delta(t)$ is a continuous function having its maximum M over (0, 1), and we have by (2.10)

$$\Delta(t) \leqslant \varepsilon \cdot c_1(y_0) + c_1(y_0)Mt.$$

Inserting again this into (2.20)

$$\Delta(t) \leqslant \varepsilon \cdot c_1(y_0) + \varepsilon c_1^2(y_0) + c_1(y_0)\frac{Mt^2}{2},$$

and proceeding similarly we get

$$\left\{\begin{aligned} &E[y_\Delta(t) - y_{\Delta'}(t)]^2 < \varepsilon c_1(y_0)\exp[c_1(y_0)] = \varepsilon \cdot c(y_0), \\ &\qquad\qquad \rho(\Delta),\quad \rho(\Delta') < \delta. \end{aligned}\right. \tag{2.11}$$

To proceed rigorously, we now introduce the measure space (Ω, P), and write as usual $y_\Delta(t) = y_\Delta(t, \omega)$, $\omega \varepsilon \Omega$. Then by (2.11)

$$\left\{\begin{aligned} &\int_0^1\int_\Omega dt\,dP[y_\Delta(t, \omega) - y_{\Delta'}(t, \omega)]^2 \leqslant \varepsilon \cdot c_2(y_0), \\ &\qquad\qquad \rho(\Delta),\quad \rho(\Delta') < \delta. \end{aligned}\right. \tag{2.12}$$

(2.11) implies that, for every t, there exists ω-measurable L^2-limit

$$\tilde{y}(t, \omega) = \underset{(\omega)}{\text{l.i.m.}}\, y_\Delta(t, \omega), \quad \rho(\Delta) \to 0,$$

whereas (2.12) means that there exists (t, ω)-measurable L^2-limit

$$y^*(t, \omega) = \underset{(t, \omega)}{\text{l.i.m.}}\, y_\Delta(t, \omega), \quad \rho(\Delta) \to 0. \tag{2.13}$$

Note that $\tilde{y}_\Delta(t, \omega)$ is not necessarily (t, ω)-measurable, but for almost all t, we have

$$\tilde{y}(t, \omega) = y^*(t, \omega) \tag{2.14}$$

for almost all $\omega \in \Omega$. Let

$$a_\Delta(t, \omega) = a[t_\nu, y_\Delta(t_\nu, \omega)], \quad t_\nu \leqslant t < t_{\nu+1}, \\ \nu = 0, 1, \ldots, n-1,$$

and similarly with b. Then we can write

$$y_\Delta(t) = y_0 + \int_0^t a_\Delta(\tau, \omega)\, d\tau + \int_0^t b_\Delta(\tau, \omega)\, dx(\tau). \tag{2.15}$$

Making use of an inequality similar to (2.7) we can show that

$$\int \{a_\Delta(t, \omega) - a[t, \tilde{y}(t, \omega)]\}^2 dP < \varepsilon c(y_0), \\ \varepsilon \to 0, \quad \text{as} \quad \rho(\Delta) \to 0.$$

Hence

$$\int_0^1 \int_\Omega \{a_\Delta(t, \omega) - a[t, y^*(t, \omega)]\}^2 dt\, dP \\ = \int_0^1 dt \int_\Omega \{a_\Delta(t, \omega) - a[t, \tilde{y}(t, \omega)]\}^2 dP \to 0, \quad \rho(\Delta) \to 0,$$

uniformly in y_0 belonging to every finite y_0-interval. We have seen in the above that $y^*(t, \omega)$, $a(t, y^*(t, \omega))$ are the (t, ω)-limit of step-functions, in other words they belong to the classe $\bar{S}$ considered by Itō [9]. Therefore, on making $\rho(\Delta) \to 0$ in (2.15), we have for almost all $\omega \in \Omega$

$$\tilde{y}(t, \omega) = y_0 + \int_0^t a[\tau, y^*(\tau)]\, d\tau + \int_0^t b[\tau, y^*(\tau)]\, dx(\tau).$$

Hence $\tilde{y}(t, \omega)$ has the «continuous kernel» (see [9], [11]) $y(t, \omega) = y(t)$ and we can write

$$y(t) = y_0 + \int_0^t a[\tau, y^*(\tau)]\, d\tau + \int_0^t b[\tau, y^*(\tau)]\, dx(\tau) \tag{2.16}$$

for almost all $\omega \in \Omega$. $y(t, \omega)$ is (t, ω)-measurable and, for every t fixed,

(2.17) $$y(t, \omega) = \tilde{y}(t, \omega)$$

for almost all ω. Combination of (2.14) and (2,17) gives

$$y(t, \omega) = y^*(t, \omega)$$

for almost all $(t, \omega) \in [0, 1] \times \Omega$. Hence we can insert $y(t, \omega)$ into the right-hand member of (2.16) and write

$$y(t) = y_0 + \int_0^t a(\tau, y(\tau))d\tau + \int_0^t b(\tau, y(\tau))dx(\tau).$$

This completes the proof.

3. *Convergence of a sequence of Markov processes.* In this section we shall deal with a convergence problem of the probability laws of a sequence of Markov processes. Bernstein's paper is concerned with the same problem. For practical use, the following formulation, in which Bernstein's formulation is contained, seems sometimes convenient (c. f. [13], pp. 24-59).

THEOREM 2. *Let $y^{(n)}(t)$, $0 \leqslant t \leqslant 1$, be a sequence of Markov processes whose transition probabilities satisfy* $(y^{(n)}(0) = y_0)$

(C_1') $$\int_{|y-x|>\delta(1+x^2)^{1/2}} dF^{(n)}(s, x; s+\Delta s, y) = \eta(\Delta s, n, x)\Delta s,$$

(C_2') $$\int_{|y-x|\leq\delta(1+x^2)^{1/2}} (y-x)dF^{(n)}(s, x; s+\Delta s, y) = a(s, x)\Delta s + \varepsilon(\Delta s, n, x)(1+x^2)^{1/2}\Delta s,$$

(C_3') $$\int_{|y-x|\leq\delta(1+x^2)^{1/2}} (y-x)^2 dF^{(n)}(s, x; s+\Delta s, y) = b^2(s, x)\Delta s + \varepsilon(\Delta s, n, x)(1+x^2)\Delta s,$$

where the functions a, b, ε, η satisfy the conditions stated in the following.

(A) *$a(s, x)$, $b(s, x)$ are continuous in (s, x), satisfy* (1.1), *continuously differentiable twice with respect to x and, as $|x| \to \infty$, $|a_{xx}| + |b_{xx}| = O(|x|^k)$ for certain $k > 0$.*

(B) *When an arbitrary $\delta > 0$ is fixed,* (i) *ε, η are bounded*

(3.1) $$|\varepsilon(\Delta s, n, x)| + |\eta(\Delta s, n, x)| \leqslant M,$$

M being independent of s, Δs, n, x; (ii)

(3.2) $$|\eta(\Delta s, n, x)| \leqslant \eta(\Delta s, n), \lim_{\Delta s \to 0} \overline{\lim_{n\to\infty}} \eta(\Delta s, n) = 0,$$

(3.3) $$|\varepsilon(\Delta s, n, x)| \leqslant \varepsilon(\Delta s, n)c(x), \lim_{\Delta s \to 0} \overline{\lim_{n\to\infty}} \varepsilon(\Delta s, n) = 0.$$

Then $y^{(n)}(t)$ *converges in probability law to the solution* $y(t)$ *of* (1.2) *i.e., for any set of t-values,* $0 \leqslant t_0 \leqslant t_1 \leqslant \cdots \leqslant t_k \leqslant 1$, *and* $a_\nu \leqslant b_\nu$, $\nu = 0, 1, \ldots, k$, *we have*

$$\text{Pr. } \{a_\nu \leqslant y^{(n)}(t_\nu) \leqslant b_\nu,\ \nu = 0, 1, \ldots, k\} \to \text{Pr. } \{a_\nu \leqslant y(t_\nu) \leqslant b_\nu,\ \nu = 0, 1, \ldots, k\},\ n \to \infty.$$

The regularity conditions imposed on a, b are weaker than those in Bernstein [1]. But they are expected to be weakened still more, in comparison with the corresponding conditions in Theorem 7. In the temporally homogeneous case, a similar problem was treated in Khintchine [13]. The formulations adopted there are not sufficient to apply to the limit theorem due to Kolmogorov-Smirnov (see § 3).

To prove the theorem we require several lemmas. The first one is Lemma 2.

LEMMA 2. *Let*

$$(3.4) \qquad z = z(y) = \int_0^y \frac{dy}{(1 + y^2)^{1/2}} = \log(y + \sqrt{1 + y^2}), \quad y = \sinh z$$

and write

$$z^{(n)}(t) = z[y^{(n)}(t)], \quad G^{(n)}(s, x;\ t, y) = \text{Pr. } \{z^{(n)}(t) \leqslant y \mid z^{(n)}(s) = x\}$$

then $z^{(n)}(t)$ *satisfies the conditions*

$$(C_1'') \quad \int_{|y-x| \geq \delta} dG^{(n)}(s, x;\ s + \Delta s, y) = \eta(\Delta s, n, x)\Delta s,$$

$$(C_2'') \quad \int_{|y-x| \leq \delta} (y - x)\, dG^{(n)}(s, x;\ s + \Delta s, y) = A(s, x)\Delta s + \varepsilon(\Delta s, n, x)\Delta s,$$

$$(C_3'') \quad \int_{|y-x| \leq \delta} (y - x)^2\, dG^{(n)}(s, x;\ s + \Delta s, y) = B^2(s, x)\Delta s + \varepsilon(\Delta s, n, x)\Delta s,$$

where

$$A(s, x) = \frac{a(s, \sinh x)}{\cosh x} - \frac{1}{2}\,\frac{\sinh x}{\cosh^3 x}\, b^2(s, \sinh x),$$

$$B(s, x) = \frac{b(s, \sinh x)}{\cosh x}.$$

The transformation (3.4) is due to Bernstein [1]. The conditions $(C_1'') - (C_3'')$ are in the same form as $(C_1') - (C_3')$, except for the factor $1 + x^2$. Another important point is that $A(s, x)$, $B(s, x)$ are now uniformly bounded

$$|A(s, x)| + |B(s, x)| \leqslant M,$$

with M independent of s, x, and satisfy the conditions in (A) of Theorem 2. We shall refere to this situation as the uniform case. Reduction to the uniform case will be used to simplify considerations with no restriction on generality.

Proof of Lemma 2. Since (3.4) is a $1-1$ mapping, $z^{(n)}(t)$ is also a Markov process. Given y_0, let us write $z_0=z(y_0)$, then to an interval $z_0-\delta\leqq z\leqq z_0+\delta$, $\delta>0$, there corresponds a y-interval $I(z_0, \delta)$. When δ is small and $|y_0|$ is large $I(z_0, \delta)$ can be approximately written $|y-y_0|\leqq\rho(y_0, \delta)$, $\rho(y_0, \delta)\sim c\delta(1+y_0^2)^{1/2}$. Hence

$$(3.5)\quad \int_{|z-z_0|>\delta} dG^{(n)}(s, z_0;\ s+\Delta s, z)\leqq\int_{|y-y_0|>c\delta(1+y_0^2)^{1/2}} dF^{(n)}(s, y_0;\ s+\Delta s, y)$$

which proves (C_1''). Next

$$(3.6)\quad \begin{cases} \displaystyle\int_{|z-z_0|\leq\delta}(z-z_0)\,dG^{(n)}(s, z_0;\ s+\Delta s, z) \\ \displaystyle =\int_{|y-y_0|\leq c\delta(1+y_0^2)^{1/2}}[z(y)-z(y_0)]\,dF^{(n)}(s, y_0;\ s+\Delta s, y] \\ \displaystyle =\frac{1}{\sqrt{1+y_0^2}}[a(s, y_0)\Delta s+\varepsilon(\Delta s, n, y_0)(1+y_0^2)^{1/2}\Delta s] \\ \displaystyle -\frac{y_0}{2(1+y_0^2)^{3/2}}[b^2(s, y_0)\Delta s+\varepsilon(\Delta s, n, y_0)(1+y_0^2)^{1/2}\Delta s] \\ \displaystyle +\int_{|y-y_0|\leq\delta(1+y_0^2)^{1/2}}\frac{c'|y-y_0|^3}{\{1+[y+\theta(y-y_0)]^2\}^{3/2}}\,dF^{(n)},\quad 0<\theta<1. \end{cases}$$

The last integral can be estimated, by (C_1'), (C_3'), as follows,

$$(3.7)\quad \begin{cases} \displaystyle\int_{|y-y_0|\leq\delta(1+y_0^2)^{1/2}}\frac{|y-y_0|^3}{\{1+[y+\theta(y-y_0)]^2\}^{3/2}}\,dF^{(n)} \\ \displaystyle =\left\{\int_{|y-y_0|\leq\varepsilon(1+y_0^2)^{1/2}}+\int_{\varepsilon(1+y_0^2)^{1/2}\leq|y-y_0|\leq\delta(1+y_0^2)^{1/2}}\right\}dF^{(n)} \\ \displaystyle \leqq\frac{c}{(1+y_0^2)^{3/2}}\left\{\varepsilon(1+y_0^2)^{1/2}\int_{|y-y_0|\leq\delta(1+y_0^2)^{1/2}}(y-y_0)^2\,dF^{(n)}\right. \\ \displaystyle \qquad\qquad\left.+\ \delta(1+y_0^2)^{3/2}\int_{\varepsilon(1+y_0^2)^{1/2}\leq|y-y_0|}dF^{(n)}\right\} \\ \displaystyle =\frac{c\varepsilon}{1+y_0^2}\{b^2(s, y_0)+\varepsilon(\Delta s, n, y_0)(1+y_0^2)\Delta s\}+c\delta\eta_1(\Delta s, n, y_0), \end{cases}$$

η_1 having the same property as η. The right-hand side of (3.7) can easily be seen to be equal to $\varepsilon_1(\Delta s, n, y_0)$, where $\varepsilon_1(\Delta s, n, x)$ has the same behavior as $\varepsilon(\Delta s, n, x)$. Substituting this estimate into (3.6) we get (C_2''). In the same way, using the Taylor expansion of $z(y) - z(y_0)$ up to the remainder term of fourth power of $y - y_0$, we get (C_3'').

LEMMA 3. *Let $x(t)$ be the Brownian motion considered in § 2, y_0 a real parameter, and $f(y_0, t, \omega)$ a random function such that $f(y_0, t, \omega) \in \bar{S}$ (for the class $\bar{S}$ see the last paragraph in the proof of Theorem 1.), and to divisions $\Delta = \Delta(t_0, t_1, \ldots, t_n)$ of (0, 1) there correspond step-functions $f_\Delta(y_0, t, \omega) \in S$ and hold*

(i) $$\int_\Omega [f_\Delta(y_0, t, \omega) - f(y_0, t, \omega)]^2 dP < \varepsilon \cdot c(y_0),$$
$$\text{with } \varepsilon \to 0, \text{ as } \rho(\Delta) \to 0,$$

(ii) $$\int_\Omega f_\Delta^2(y_0, t, \omega) dP < c(y_0).$$

Then we have, uniformly in y_0 belonging to every finite interval, in the L^2-sense

$$\underset{\varrho(\Delta)\to 0}{\text{l.i.m.}} \sum\nolimits_{(t)} f_\Delta(y_0, t, \omega) \psi(\underset{(\Delta\tau_\nu)}{\Delta x_\nu}) = \int_0^t f(y_0, \tau, \omega) \, d\underset{(d\tau)}{x(\tau)},$$

where $\psi(x)$ is an odd function such that

$$\psi(x) = x \qquad |x| \leqslant \lambda,$$
$$= 0 \qquad |x| > \lambda,$$

$\lambda > 0$ being an arbitrary fixed number.

Proof. We have

$$\int \left\{ \int_0^t f(y_0, \tau, \omega) dx(\tau) - \sum\nolimits_{(t)} f_\Delta(y_0, t_\nu, \omega) \psi(\Delta x_\nu) \right\}^2 dP$$

$$\leqslant 2 \int \left\{ \int_0^t [f(y_0, \tau, \omega) - f_\Delta(y_0, \tau, \omega)] dx(\tau) \right\}^2 dP$$

$$+ 2 \int \left\{ \sum\nolimits_{(t)} f_\Delta(y_0, t_\nu, \omega) [\Delta x_\nu - \psi(\Delta x_\nu)] \right\}^2$$

$$\leqslant 2 \int_0^1 \int_\Omega (f_\Delta - f)^2 d\tau dP + c \sum_0^{n-1} \left(\int f_\Delta^2 dP \right) (\Delta t_\nu)^2 \leqslant \eta \cdot c(y_0)$$

with $\eta \to 0$, as $\rho(\Delta) \to 0$. This proves the first half of the lemma. The second half is proved in the same way.

In what follows y_0 is a real parameter, st. lim means the limit in probability, and convergence with uniformity (y_0) means the uniform convergence in y_0 belonging to every finite interval.

LEMMA 4. *Let $x_n(y_0)$ be a sequence of random variables depending on a parameter y_0 such that*

$$E(|x_n(y_0)|^\alpha) < c(y_0) \qquad (\alpha > 1)$$

end there exists

$$\text{st. lim}\, x_n(y_0) = x(y_0)$$

uniformly (y_0), i.e. the Fréchet distance $r[x_n(y_0),\ x(y_0)] < \varepsilon c(y_0)$, with $\varepsilon \to 0$, as $n \to \infty$. Then, for every β, $1 < \beta < \alpha$, there exists in the L^β-sense

$$\underset{n\to\infty}{\text{l.i.m.}}\, x_n(y_0) = x(y_0)$$

uniformly (y_0).

LEMMA 5. *Let $f(x)$, $-\infty < x < \infty$, be any continuous function, and $x_n(y_0)$ be a sequence of random variables such that*

$$E(|x_n(y_0)|) < c(y_0)$$

and there exists

$$\text{st. lim}\, x_n(y_0) < x(y_0)$$

uniformly (y_0).

Then

$$\text{st. lim} f[x_n(y_0)] = f[x(y_0)]$$

uniformly (y_0).

The proofs of these lemmas are easy, and may be omitted.

LEMMA 6. *If $a(s, x)$, $b(s, x)$ satisfy the conditions under Theorem 2 and $y(t)$ is the solution of (1.2), then there exist $\partial E(\exp[izy(t)])/\partial y_0$, $\partial^2 E(\exp[izy(t)])/\partial y_0^2$. They are continuous in y_0, and written as*

$$\frac{\partial}{\partial y_0} E(\exp[izy(t)]) = izE(\exp[\Psi(t) + izy(t)]), \tag{3.8}$$

$$\begin{cases} \dfrac{\partial^2}{\partial y_0^2} E(\exp[izy(t)]) = E(\exp[izy(t) + \Psi(t)]\{iz\displaystyle\int_0^t \exp[\Psi(\tau)] \\ \qquad \times [(a'' - b'b'')d\tau + b''dx(\tau)] - z^2\exp[\Psi(t)]\}), \end{cases} \tag{3.9}$$

where

$$\Psi(t) = \int_0^t \left(a' - \frac{b'^2}{2}\right)d\tau + \int_0^t b'\,dx(\tau),$$

$$a' = a_y[\tau,\ y(\tau)], \qquad a'' = a_{yy}[\tau,\ y(\tau)] \quad \text{etc.}$$

To prove Lemma 6 we require the

LEMMA 7. *Let $f(t, y)$ be a continuous function of $t, y, 0 \leqslant t \leqslant 1, -\infty < y < \infty$, and $x_n(y_0)$ the random variable satisfying the same conditions as in Lemma 5. Then*

$$\text{st. lim } f[s, x_n(y_0)] = f[t, x(y_0)]$$

as $s \to t$, $n \to \infty$, uniformly in t and y_0. In particular if $|f(x)| < c(1 + |x|^p)$ for certain $p > 0$, and if y_ν are the variables in Theorem 1, then for $k \geqslant 1$, $t_\nu \leqslant t < t_{\nu+1}$, we have, uniformly in t and y_0,

$$E\{f_\Delta(t) - f[t, y(t)]\}^{2k} \to 0, \qquad \rho(\Delta) \to 0,$$

where

$$f_\Delta(t) = f(t_\nu, y_\nu), \qquad t_\nu \leqslant t < t_{\nu+1}.$$

Proof. If we consider the combined variable (t, y_0) as a parameter, the first half of the lemma is a direct consequence of Lemma 5. The second half then follows from the first half. To see this we have only to note that $E|f(t_\nu, y_\nu)|^\beta \leqslant c(1 + |y_\nu|^{p\beta}) \leqslant c(y_0)$, and apply Lemma 4.

Proof of Lemma 6. Let $\Delta = \Delta(s_0, s_1, \ldots, s_m)$ be a division of $(0, t)$, $0 = s_0 < s_1 < \cdots < s_m = t$, and consider the difference equations

$$\text{(3.10)} \qquad \begin{cases} y_1 = y_0 + a(t_0, y_0)\Delta t_0 + b(t_0, y_0)\psi(\Delta x_0), \\ y_2 = y_1 + a(t_1, y_1)\,\Delta t_1 + b(t_1, y_1)\psi(\Delta x_1), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_m = y_{m-1} + a(t_{m-1}, y_{m-1})\Delta t_{m-1} + b(t_{m-1}, y_{m-1})\psi(\Delta x_{m-1}), \end{cases}$$

which is obtained from (2.1) replacing Δx_ν by $\psi(\Delta x_\nu)$. Since the truncated randon variables $\psi(\Delta x_\nu)$ are very good approximations, when $\rho(\Delta) \to 0$, to Δx_ν, we can easily show that y_m are also good approximations to $y_\Delta(t)$ and, in the L^2-sense, $\underset{\rho(\Delta)\to 0}{\text{l. i. m.}}\, y_m = y(t)$. Hence Lemma 7 also holds for the y_ν in (3.10), instead for those in Theorem 1.

We have

$$\text{(3.11)} \qquad \frac{\partial}{\partial y_0} E(e^{izy_m}) = izE\left(e^{izy_m}\frac{\partial y_m}{\partial y_0}\right),$$

$$\text{(3.12)} \qquad \frac{\partial^2}{\partial y_0^2} E(e^{izy_m}) = E\left\{e^{izy_m}\left(iz\frac{\partial y_m}{\partial y_0}\right)^2 + e^{izy_m} iz \frac{\partial^2 y_m}{\partial y_0^2}\right\},$$

differentiations under expectation signs being easily justified. But differentiation and successive substitutions in (3.10) yield

$$u_m = \frac{\partial y_m}{\partial y_0} = \prod_{\nu=0}^{m-1} A_\nu, \quad A_\nu = 1 + a'(t_\nu, y_\nu)\Delta t_\nu + b'(t_\nu, y_\nu)\psi(\Delta x_\nu).$$

Also differentiating twice (3.10)

$$v_m = \frac{\partial^2 y_m}{\partial y_0^2} = v_{m-1} A_{m-1} + u_{m-1}^2 B_{m-1}, \quad (u_0 \equiv 1,\ v_0 \equiv 0),$$

$$B_\nu = a''(t_\nu,\ y_\nu)\Delta t_\nu + b''(t_\nu,\ y_\nu)\psi(\Delta x_\nu),$$

and hence by successive substitutions

$$v_m = u_m \sum_{\nu=0}^{m-1} \frac{u_\nu B_\nu}{A_\nu}. \tag{3.13}$$

Remembering that a', b' are bounded by (1.1) we can suppose now

$$\operatorname*{Max}_{t,\,y} |a'(t,\ y)|\rho(\Delta) + \operatorname*{Max}_{t,\,y} |b'(t,\ y)|\varepsilon < \mu,$$

μ being able to be made as small as we please, if only $\rho(\Delta)$ and ε are sufficiently small. Therefore

$$\lambda_m = E(|u_m|) = E(u_m)$$
$$= E\left\{\prod_{\nu=0}^{m-2} A_\nu [1 + a'(t_{m-1},\ y_{m-1})\Delta t_{m-1}]\right\} \leqq \lambda_{m-1} + c\lambda_{m-1}\Delta t_{m-1}.$$

Hence

$$E(|u_m|) \leqq c_1.$$

In the same way $E(u_m^2) \leqq c_2$, and in general we can show that

$$E(u_m^{2p}) \leqq c_{2p}, \tag{3.14}$$

c_{2p} being independent of y_0, t.

We can write

$$u_m = \exp\left[\sum_{\nu=0}^{m-1} \log A_\nu\right] = \exp\left[\sum_{0}^{m-1}\left(a'_\nu \Delta t_\nu + b'_\nu \psi_\nu - \frac{b_\nu'^2}{2}\Delta t_\nu\right) + R_\Delta\right] \tag{3.15}$$

where

$$\left\{\begin{aligned} R_\Delta &= \sum_{0}^{m-1}\left\{O(\Delta t_\nu)^2 + O(\Delta t_\nu |\psi_\nu|) + O(|\psi_\nu|^3) + \frac{1}{2} b_\nu'^2(\Delta t_\nu - \psi_\nu^2)\right\}, \\ & a'_\nu = a'(t_\nu,\ y_\nu), \quad b'_\nu = b'(t_\nu,\ y_\nu), \quad \psi_\nu = \psi(\Delta x_\nu). \end{aligned}\right. \tag{3.16}$$

O's in (3.16), all of which are products of a'_ν, b'_ν, are uniformly bounded. Also from (3.13)

$$(3.17)\quad \begin{cases} v_m = u_m \sum_0^{m-1} u_\nu \{(a''_\nu - b'_\nu b''_\nu)\Delta t_\nu + b''_\nu \psi_\nu\} + S_\Delta, \\ S_\Delta = u_m \sum_0^{m-1} u_\nu w_\nu, \quad b''_\nu = b''(t_\nu, y_\nu), \end{cases}$$

$$(3.18)\quad w_\nu = O(\Delta t_\nu)^2 + O(|\psi_\nu|^3) + O(|\Delta t_\nu \psi_\nu|) + b'_\nu b''_\nu (\Delta t_\nu - \psi_\nu^2).$$

O's in (3.18), all of which are products of a'_ν, b'_ν, a''_ν, b''_ν can be dominated by $c(1 + y_\nu^{2p})$ for a certain $p > 0$.

To prove convergence of u_m, v_m, we first observe that by the Gaussian distribution of $\Delta x = x(t + \Delta t) - x(t)$

$$\begin{aligned} E|\Delta t - \psi^2(\Delta x)|^k &\leqslant c_1 E|\Delta t - (\Delta x)^2|^k + c_2 E|\psi^2(\Delta x) - (\Delta x)^2|^k \\ &= c_1(2\pi\Delta t)^{-1/2}\int |\Delta t - u^2|^k e^{-u^2/2\Delta t} du \\ &+ 2c_2(2\pi\Delta t)^{-1/2}\int_\varepsilon^\infty u^{2k} e^{-u^2/2\Delta t} du = O(\Delta t)^k, \quad k > 0. \end{aligned}$$

Also

$$(3.19)\quad E(\Delta t - \psi^2(\Delta x)) = O(\Delta t)^2, \quad E(|\psi(\Delta x)|^k) = O(\Delta t)^{k/2}.$$

Substituting (3.19) into (3.18) we get

$$(3.20)\quad E(w_\nu) = O(\Delta t)^{3/2}, \quad E(|w_\nu|^k) = O(\Delta t)^k.$$

Now let us write

$$(3.21)\quad \left(\sum u_\nu w_\nu\right)^4 = \sum_{p+q+r+s=4} \frac{4!}{p!\,q!\,r!\,s!}(u_\nu w_\nu)^p (u_{\nu'} w_{\nu'})^q (u_{\nu''} w_{\nu''})^r (u_{\nu'''} w_{\nu'''})^s$$

and consider the mean value, then we encounter, e.g. a term of the form $m = E(u_\nu^2 w_\nu^2 u_{\nu'} w_{\nu'} u_{\nu''} w_{\nu''})$ $(\nu < \nu' < \nu'')$. But (3.20) and Hölder's inequality give

$$\begin{aligned} |m| &\leqslant cE|u_\nu^2 w_\nu^2 u_{\nu'} w_{\nu'} u_{\nu''}|\,\Delta t_{\nu''}^{3/2} \\ &\leqslant c(E|u_\nu^6 u_{\nu'}^3 u_{\nu''}^3|)^{1/3}(E w_\nu^6)^{1/3}(E w_{\nu'}^3)^{1/3}\Delta t_{\nu''}^{3/2} \\ &\leqslant c'\Delta t_\nu^2 \Delta t_{\nu'} \Delta t_{\nu''}^{3/2}. \end{aligned}$$

In evaluation of the mean value of (3.21) this contributes

$$c' \sum_{\nu, \nu', \nu''} \Delta t_\nu^2 \Delta t_{\nu'} \Delta t_{\nu''}^{3/2} \leqslant c'\rho(\Delta)^{3/2}$$

and similarly for other terms, having

$$E\Big(\sum_{\nu} u_\nu w_\nu\Big)^4 \leqslant c\,\rho(\Delta)^{1/2} \rightarrow 0 \tag{3.22}$$

uniformly (y_0). Hence by (3.14) and (3.22)

$$E(S_\Delta^2) \rightarrow 0, \tag{3.23}$$

and similarly

$$E(R_\Delta^2) \rightarrow 0. \tag{3.24}$$

either uniformly (y_0).

Substitute (3.24) into (3.15), then by Lemma 3 and Lemma 7 we obtain in the L^2-sense

$$\text{l. i. m.} \sum_0^{m-1} \log A_\nu = \Psi(t)$$

uniformly in y_0, t. Hence by Lemma 4, (3.14) and Lemma 5 with $f(x) = \exp(x)$, for every $\beta \geqslant 1$, we have

$$\underset{\rho(\Delta)\rightarrow 0}{\text{l. i. m.}}\, u_m = \exp[\Psi(t)] \tag{3.25}$$

in the L^β-sense, uniformly (y_0). This together with (3.11) gives (3.8) and its continuity in y_0.

To prove (3.9), first we observe, in the same way as in (3.22), that

$$E\Big\{\sum u_\nu [(a_\nu'' - b_\nu b_\nu'')\Delta t_\nu + b_\nu'' \psi_\nu]\Big\}^4 \leqslant c_1(y_0), \tag{3.26}$$

and so

$$E(v_m^2) \leqslant c_2(y_0). \tag{3.27}$$

Next by (3.25) and Lemma 7, for $t_\nu \leqslant t < t_{\nu+1}$, we get in the L^2-sense, as $\rho(\Delta) \rightarrow 0$

$$u_\nu(a_\nu'' - b_\nu' b_\nu'') \rightarrow \exp[\Psi(t)]\{a''[t, y(t)] - b'[t, y(t)]\, b''[t, y(t)]\},$$

$$u_\nu b_\nu'' \rightarrow \exp[\Psi(t)]\, b''(t, y)$$

uniformly in y_0 and t.

Hence again by (3.25),

$$v_m \rightarrow \Psi(t) \int_0^t \Psi(t)\{(a'' - b' b'')\, d\tau + b''\, dx(\tau)\} \tag{3.28}$$

at least in the L^1-sense, uniformly (y_0). Substituting (3.25) and (3.28) into (3.12) we obtain (3.9) and its continuity in y_0. This completes the proof of Lemma 6.

LEMMA 8. *Let* $\xi_0, \xi_1, \ldots$ *be random variables* (not necessarily independent each other) *such that*

$$E(\xi_i | \xi_1, \xi_2, \ldots, \xi_{i-1}) = 0, \quad \beta_i = E(\xi_i^2) < \infty.$$

Let

$$B_n = \sum_0^n \beta_i, \quad S_\nu = \sum_0^\nu \xi_i,$$

then

$$\Pr.\{\operatorname*{Max}_{1 \leq \nu \leq n} |S_\nu| > t\sqrt{B_n}\} \leqq t^{-2}.$$

This is a well-known generalization of Kolmogorov's inequality. For the proof see [1].

LEMMA 9. *Let* $\sum = (s_1, s_2, \ldots, s_N)$ *be an arbitrary number of points in* (0, 1). *Then under the assumptions of Theorem* 2 *we have*

$$\lim_{K \to \infty} \overline{\lim_{n \to \infty}} \Pr.\{\operatorname*{Max}_{1 \leq \nu \leq N} |y^{(n)}(s_\nu)| > K\} = 0,$$

the limit in K being uniform in the choice of $\sum$.

Proof. Let $\Delta(t_0, t_1, \ldots, t_m)$ be a division of (0, 1), of which $\sum$ is a subset of the division points and $\rho(\Delta)$ is sufficiently small, and note that according to Lemma 2, we have only to prove the uniform case of the lemma. Let

$$\begin{aligned} \varphi_\delta(x) &= x \quad \text{for} \quad |x| \leqq \delta \\ &= 0 \quad \text{otherwise} \end{aligned}$$

and let

$$y_\nu = y^{(n)}(t_\nu).$$

Then, by (C_2'') and boundedness of a, b, if we put

$$c_{\nu-1} = E(\varphi_\delta(y_\nu - y_{\nu-1}) | y_{\nu-1}) = \int_{|y_\nu - y_{\nu-1}| < \delta} (y_\nu - y_{\nu-1}) \, dF^{(n)},$$

we get

(3.29) $$|c_{\nu-1}| \leqq c \, \Delta t_{\nu-1}.$$

Also, if we write

$$\xi_{\nu-1} = \varphi_\delta(y_\nu - y_{\nu-1}) - c_{\nu-1},$$

then $\xi_0, \xi_1, \ldots,$ satisfy the conditions of Lemma 8 and (C_3'') gives

$$(3.30)\quad \begin{cases} \beta_{\nu-1}=E(\xi_{\nu-1}^2)\leqq 2\Big[\int_{-\infty}^{\infty} dF^{(n)}(t_0, y_0; t_{\nu-1}, y_{\nu-1})\int_{|y_\nu-y_{\nu-1}|\leqq\delta}(y_\nu-y_{\nu-1})^2 \\ \times dF^{(n)}(t_{\nu-1}, t_\nu; y_{\nu-1}, y_\nu)+E(c_{\nu-1}^2)\Big]\leqq 2[c_1\Delta t_{\nu-1}+c^2(\Delta t_{\nu-1})^2]\leqq c_2\Delta t_{\nu-1}. \end{cases}$$

Now

$$\Pr.\{\operatorname*{Max}_{1\leqq\nu\leqq n}|y_\nu|>K\}\leqq\sum_{\nu=1}^{m}\Pr.\{|y_\nu-y_{\nu-1}|>\delta\}$$

$$+\Pr.\{\operatorname*{Max}_{1\leqq\nu\leqq n}|y_\nu-y_{\nu-1}|\leqq\delta,\quad \operatorname*{Max}_{1\leqq\nu\leqq n}|y_0+\sum_{\mu=1}^{\nu}\varphi_\delta(y_\mu-y_{\mu-1})|>K\}.$$

But by (C_1'')

$$\Pr.\{|y_\nu-y_{\nu-1}|>\delta\}\leqq\int_{-\infty}^{\infty}dF^{(n)}(t_0, y_0;\ t_{\nu-1},\ y_{\nu-1})\int_{|y_\nu-y_{\nu-1}|>\delta}dF^{(n)}(t_{\nu-1}, y_{\nu-1}; t_\nu, y_\nu)$$
$$\leqq\eta(\Delta t_{\nu-1},\ n)\Delta t_{\nu-1},$$

and by (3.29), (3.30)

$$\Pr.\{\operatorname*{Max}_{\nu}|y_0+\sum_1^\nu\varphi_\delta(y_\mu-y_{\mu-1})|>K\}$$

$$\leqq\Pr.\{\operatorname*{Max}_{\nu}\Big|\sum_1^\nu\xi_{\mu-1}\Big|>K-|y_0|-\sum_1^m|c_{\mu-1}|\}$$

$$\leqq\frac{c_2}{(K-|y_0|-c)^2}.$$

Hence

$$\overline{\lim_{n\to\infty}}\Pr.\{\operatorname*{Max}_{1\leqq\nu\leqq N}|y^{(n)}(t_\nu)|>K\}\leqq\overline{\lim_{n\to\infty}}\Pr.\{\operatorname*{Max}_{1\leqq\nu\leqq m}|y^{(n)}(t_\nu)|>K\}$$

$$\leqq\overline{\lim_{n\to\infty}}\sum_{\nu=1}^{m}\eta(\Delta t_{\nu-1},\ n)\Delta t_{\nu-1}+\frac{c_2}{(K-|y_0|-c)^2}$$

$$\leqq\eta+\frac{c_2}{(K-|y_0|-c)^2},$$

which can be made as small as we please, if $\rho(\Delta)$ is small and K is large.

Proof of Theorem 2. By Lemma 2, we have only to prove the uniform case. Indeed, if we use the transformation $z(t)=z(y(t))$, $z(t)$ satisfies the Itō equation (1.2) with A, B instead of a, b, i. e. $z(t)$ is the unique solution in the uniform case (see Lemma 11), and convergence relation between $z^{(n)}(t)$ and $z(t)$ is equivalent to the corresponding one between $y^{(n)}(t)$ and $y(t)$. To simplify notations we shall prove the simplest case $k=1$, the following method being applicable to the general case as well.

Consider a division $\Delta(t_0, t_1, \ldots, t_m)$, and put $y_\nu = y(t_\nu)$. Introduce a monotone function Φ such that $\Phi(x)$ is continuously differentiable twice and

$$\begin{aligned}\Phi(x) &= x \qquad\qquad \text{for } |x| \leqslant K \\ &= \pm(K+1) \quad \text{for } |x| > K+1,\end{aligned}$$

where K is a large fixed number. Now define new variables $\tilde{y}_\nu$, $\nu = 0, 1, 2, \ldots, m-1$, whose transition probability is given by

$$\operatorname{Pr.}\{\tilde{y}_\nu \leqslant u\} = \int_{-\infty}^{\infty} F(t_{\nu-1}\Phi(y_\nu);\ 1,\ u)\, dF^{(n)}(t_0, y_0;\ t_\nu, y_\nu),$$
$$\nu = 1, 2, \ldots, m-1,$$

where $F(s, x;\ t, y)$ is the transition probability of $y(t)$, and let

$$\operatorname{Pr.}\{\tilde{y}_0 \leqslant u\} \equiv F(t_0, \Phi(y_0);\ 1,\ u).$$

If we write

$$f(s, x) = \int_{-\infty}^{\infty} e^{izy}\, dF(s, x;\ 1,\ y)$$

we have

$$E\{e^{i\tilde{y}_{\nu-1}z}\} = \int_{-\infty}^{\infty} dF^{(n)}(t_0, y_0;\ t_{\nu-1}, y_{\nu-1}) \int_{-\infty}^{\infty} f(t_\nu, y_\nu)\, dF(t_{\nu-1}, \Phi(t_{\nu-1});\ t_\nu, y_\nu), \tag{3.31}$$

$$E\{e^{i\tilde{y}_\nu z}\} = \int_{-\infty}^{\infty} dF^{(n)}(t_0, y_0;\ t_{\nu-1}, y_{\nu-1}) \int_{-\infty}^{\infty} f(t_\nu, \Phi(y_\nu))\, dF^{(n)}(t_{\nu-1}, y_{\nu-1};\ t_\nu, y_\nu). \tag{3.32}$$

By the Taylor expansion

$$\left\{\begin{aligned} f(t_\nu, y_\nu) &= f[t_\nu, \Phi(y_{\nu-1})] + f'[t_\nu, \Phi(y_{\nu-1})][y_\nu - \Phi(y_{\nu-1})] \\ &+ \frac{1}{2} f''[t_\nu, \Phi(y_{\nu-1})][y_\nu - \Phi(y_{\nu-1})]^2 \\ &+ \frac{1}{2}[y_\nu - \Phi(y_{\nu-1})]^2\{f''[t_\nu, \Phi(y_{\nu-1}) + \theta(y_\nu - \Phi(y_{\nu-1})] - f''[t_\nu, \Phi(y_{\nu-1})]\}, \end{aligned}\right. \tag{3.33}$$

$$\left\{\begin{aligned} f[t_\nu, \Phi(y_\nu)] &= f[t_\nu, \Phi(y_{\nu-1})] + (y_\nu - y_{\nu-1}) f'[t_\nu, \Phi(y_{\nu-1})]\Phi'(y_{\nu-1}) \\ &+ \frac{1}{2}(y_\nu - y_{\nu-1})^2 f''[t_\nu, \Phi(y_{\nu-1})]\Phi'^2(y_{\nu-1}) \\ &+ \frac{1}{2}(y_\nu - y_{\nu-1})^2 f'[t_\nu, \Phi(y_{\nu-1})]\Phi''(y_{\nu-1}) \\ &+ \frac{1}{2}(y_\nu - y_{\nu-1})^2\{f''[t_\nu, \Phi(y_{\nu-1} + \theta(y_\nu - y_{\nu-1}))]\Phi'^2[y_{\nu-1} + \theta(y_\nu - y_{\nu-1})] \\ &+ f'(t_\nu, \Phi[y_{\nu-1} + \theta(y_\nu - y_{\nu-1})])\Phi''[y_{\nu-1} + \theta(y_\nu - y_{\nu-1})] \\ &- f''[t_\nu, \Phi(y_{\nu-1})]\Phi'^2(y_{\nu-1}) - f'[t_\nu, \Phi(y_{\nu-1})]\Phi''(y_{\nu-1})\}, \qquad 0 < \theta < 1. \end{aligned}\right. \tag{3.34}$$

It should be noted that $f(s, x)$ is continuously differentiable twice by Lemma 6, therefore these expansions are possible.

Substitute (3.33), (3.34) into (3.31), (3.32), and use the conditions $(C_1'')-(C_3'')$ and the fact that the function $f''(t_\nu, \Phi(y_{\nu-1})+\theta(y_\nu-\Phi(y_{\nu-1})))-f''(t_\nu, \Phi(y_{\nu-1}))$ and the like are small, if only $|y_\nu-\Phi(y_{\nu-1})|$ is small, uniformly in $y_{\nu-1}$, and in performing interior integrations, divide the ranges of integration into $|y_\nu-\Phi(y_{\nu-1})|\leqslant\delta$, $|(y_\nu-\Phi(y_{\nu-1})|>\delta$ and $|y_\nu-y_{\nu-1}|\leqslant\delta$, $|y_\nu-y_{\nu-1}|>\delta$, with $\delta>0$ small. Then we obtain

$$
(3.35)\quad \begin{cases} E\{e^{iz\tilde{y}_{\nu-1}}\}=\displaystyle\int_{-\infty}^{\infty} dF^{(n)}(t_0, y_0;\ t_{\nu-1}, y_{\nu-1})\{f[t_\nu, \Phi(y_{\nu-1})] \\ \qquad +f'[t_\nu, \Phi(y_{\nu-1})]\,a[t_{\nu-1}, \Phi(y_{\nu-1})]\,\Delta t_{\nu-1}+\dfrac{1}{2}f''[t_\nu, \Phi(y_{\nu-1})] \\ \qquad b^2[t_\nu, \Phi(y_{\nu-1})]\,\Delta t_{\nu-1}\}+\eta(\Delta t_{\nu-1}, n)\,\Delta t_{\nu-1}+\varepsilon(\Delta t_{\nu-1}, n)\,\Delta t_{\nu-1} \\ \qquad +\varepsilon(\delta)\cdot\Delta t_{\nu-1}, \end{cases}
$$

$$
(3.36)\quad \begin{cases} E\{e^{iz\tilde{y}_\nu}\}=\displaystyle\int_{-\infty}^{\infty} dF^{(n)}(t_0, y_0;\ t_{\nu-1}, y_{\nu-1})\{f[t_\nu, \Phi(y_{\nu-1})] \\ \qquad +f'[t_\nu, \Phi(y_{\nu-1})]\,\Phi'(y_{\nu-1})\,a(t_{\nu-1}, y_{\nu-1})\,\Delta t_{\nu-1} \\ \qquad +\dfrac{1}{2}f''[t_\nu, \Phi(y_{\nu-1})]\,\Phi'^2(y_{\nu-1})\,b^2(t_{\nu-1}, y_{\nu-1})\,\Delta t_{\nu-1} \\ \qquad +\dfrac{1}{2}f'[t_\nu, \Phi(y_{\nu-1})]\,\Phi''(y_{\nu-1})\,b^2(t_{\nu-1}, y_{\nu-1})\,\Delta t_{\nu-1}\} \\ \qquad +\eta(\Delta t_{\nu-1}, n)\,\Delta t_{\nu-1}+\varepsilon(\Delta t_{\nu-1}, n)\,\Delta t_{\nu-1}+\varepsilon(\delta)\,\Delta t_{\nu-1}, \end{cases}
$$

where $\varepsilon(\delta)\to 0$, as $\delta\to 0$.
Recalling

$$
\begin{aligned}
\Phi'(u)&=0 \quad \text{for} \quad |u|>K+1, \\
&=1 \quad \text{for} \quad |u|\leqslant K, \\
\Phi''(u)&=0 \quad \text{for} \quad |u|\leqslant K.
\end{aligned}
$$

Subtract (3.35) from (3.36) and use Lemma 9, then we get

$$\overline{\lim_{n\to\infty}} |E(e^{iz\tilde{y}_0}) - E(e^{iz\tilde{y}_{m-1}})| \leq \overline{\lim_{n\to\infty}} \sum_{\nu=1}^{m-1} |E(e^{iz\tilde{y}_\nu}) - E(e^{iz\tilde{y}_{\nu-1}})|$$

$$\leq c \overline{\lim_{n\to\infty}} \sum_{\nu=1}^{m-1} \text{Pr.} \{|y^{(n)}(t_{\nu-1})| > K\} + \overline{\lim_{n\to\infty}} \sum_{\nu=1}^{m-1} [\varepsilon(\Delta t_{\nu-1}, n)$$

$$+ \eta(\Delta t_{\nu-1}, n)] \Delta t_{\nu-1} + \varepsilon(\delta) \sum_{1}^{m-1} \Delta t_{\nu-1} \leq \varepsilon(\delta) + \varepsilon[\rho(\Delta)] + \varepsilon(K),$$

where $\varepsilon(K)$, $\varepsilon[\rho(\Delta)]$, $\varepsilon(\delta) \to 0$, as $K \to \infty$, $\rho(\Delta) \to 0$, $\delta \to 0$ respectively. Also, if we write $\tilde{y}_{m-1} = \{\tilde{y}_{m-1} - \Phi(y^{(n)}_{m-1})\} + \Phi(y^{(n)}_{m-1})$, $y^{(n)}(1) = \{y^{(n)}(1) - y^{(n)}_{m-1}\} + y^{(n)}_{m-1}$, where $y^{(n)}_{m-1} = y^{(n)}(t_{m-1})$, and note that

$$E\{\exp[iz(\tilde{y}_{m-1} - \Phi(y^{(n)}_{m-1}))] | y^{(n)}_{m-1} \text{ is given}\},$$

$$E\{\exp[iz(y^{(n)}(1) - y^{(n)}_{m-1})] | y^{(n)}_{m-1} \text{ is given}\}$$

$$= 1 + O(1 - t_{m-1}), \quad \text{as} \quad n \to \infty,$$

and that $\tilde{y}_0$ has the same distribution as $y(1)$ for sufficiently large K, then we have $E(e^{iz\tilde{y}_{m-1}}) - E(e^{izy^{(n)}(1)}) \to 0$, as $n \to 0$, and then $\rho(\Delta) \to 0$. Hence we have

$$\overline{\lim_{n\to\infty}} |E(e^{izy(1)}) - E(e^{izy^{(n)}(1)})| = 0,$$

which completes the proof.

4. *Invariance principle in the Markovian scheme.* Let $X_1, X_2, \dots$ be equally distributed independent random variables. In many cases, for a certain class of functions $\Phi_n(x_1, x_2, \dots, x_n)$ the limit distribution of $\Phi_n(X_1, X_2, \dots, X_n)$ is free from the underlying distribution of X_i, and sometimes the relation holds

$$\lim \text{Pr.} \{\Phi_n(X_1, X_2, \dots, X_n) \leq \alpha\} = \text{Pr.} \{\Phi(x(\cdot)) \leq \alpha\},$$

where $\Phi(\cdot)$ is a functional and $x(\tau)$ is a stochastic process. Kac and Erdös [5], [6] have recently formulated limit theorems in this form, what they call the invariance principle. An important class of the principles is based on the several dimensional central limit theorem, $x(\tau)$ then being the Brownian motion. In this section we shall consider the same problem in the scheme of Markov process and show that our result is applicable to a rigorous treatment of Doob's method [3] proving the Kolmogorov-Smirnov limit theorem.

THEOREM 3. *Let $y^{(n)}(t)$ be a sequence of Markov processes satisfying the conditions under Theorem 2, $y(t)$ the corresponding limit process, and let $f(t)$, $g(t)$ be two continuous functions such that $f(t) < g(t)$, $0 \leqq t \leqq 1$, $f(0) < y_0 < g(0)$.*

Then we have

$$\left\{\begin{aligned} &\lim_{n\to\infty} \Pr.\{f(t) \leqq y^{(n)}(t) \leqq g(t), \quad 0 \leqq t \leqq 1\} \\ &= \Pr.\{f(t) \leqq y(t) \leqq g(t), \quad 0 \leqq t \leqq 1\}. \end{aligned}\right. \tag{4.1}$$

The above formulation of the convergence problem of the processes with a continuous time parameter does not restrict our situation. For, given an arbitrary sequence of processes $\{y^{(n)}(k),\ k = 0, 1, \dots\}$ we can construct the one with a continuous time parameter as follows,

$$y^{(n)}(t) \equiv y^{(n)}(k), \quad k-1 \leqq t < k, \quad k = 1, 2, \dots .$$

The proof of Theorem 3 goes along the same line as in [5], now using Theorem 2 instead of the central limit theorem.

Proof of Theorem 3. By the same reason as with the proof of Theorem 2 we are sufficient to prove the uniform case of Theorem 3. Also, for simplicity, we shall consider the problem with only one absorbing barrier, in which we are lacking $f(t)$ on both sides of (4.1).

Consider a division $\Delta(t_0, t_1, \dots, t_k)$ of (0, 1) and subdivision of each (t_i, t_{i+1}):

$$\begin{aligned} 0 = t_0 = s_1^{(1)} < \cdots < s_m^{(1)} < t_1 = s_1^{(2)} < \cdots < s_m^{(2)} < t_2 \\ = s_1^{(3)} < \cdots < s_m^{(k)} = t_k = 1. \end{aligned}$$

Denote by $A_{\nu,i}$ the simultaneous occurrence of

$$y^{(n)}(s_\nu^{(i)}) > f(s_\nu^{(i)}), \qquad y^{(n)}(\tau) \leqq f(\tau)$$

for $0 \leqq \tau < s_\nu^{(i)}$, τ ranges over only the division points, and put $p_{\nu,i} = \Pr.(A_{\nu,i})$. Then by (C_1'')

$$\begin{aligned} p_{\nu,i} &= \Pr.\{A_{\nu,i} \,|\, y^{(n)}(t_i) - y^{(n)}(s_\nu^{(i)})| > \varepsilon\} \\ &\quad + \Pr.\{A_{\nu,i} \,|\, y^{(n)}(t_i) - y^{(n)}(s_\nu^{(i)})| \leqq \varepsilon\} \\ &\leqq \eta(\Delta t_{\nu-1}, n) \Delta t_{\nu-1} \Pr.(A_{\nu,i}) \\ &\quad + \Pr.\{A_{\nu,i} \,|\, y^{(n)}(t_i) - y^{(n)}(s_\nu^{(i)})| \leqq \varepsilon\}, \end{aligned}$$

whence

$$\sum_{\nu,i} p_{\nu,i} = 1 - \Pr.\{y^{(n)}(s_\nu^{(i)}) \leqslant f(s_\nu^{(i)}),\ \nu = 1, \ldots, m,\ i = 1, \ldots, k\}$$

$$\leqslant \sum \eta(\Delta t_{\nu-1},\ n)\Delta t_{\nu-1} + \Pr.\{\text{at least one occurs among}$$

$$|y^{(n)}(t_i) - y^{(n)}(s_\nu^{(i)})| \leqslant \varepsilon,\ A_{\nu,i},\ \nu = 1, \ldots, m,\ i = 1, \ldots, k\}$$

$$\leqslant \sum \eta(\Delta t_{\nu-1},\ n)\Delta t_{\nu-1} + 1 - \Pr.\{y^{(n)}(t_i) \leqslant f(t_i) - 2\varepsilon,\ i = 1, \ldots, k\},$$

if $\rho(\Delta) = \operatorname{Max}(t_i - t_{i-1})$ is small. Or

$$(4.2)\quad \begin{cases} \Pr.\{y^{(n)}(t_i) \leqslant f(t_i) - 2\varepsilon,\ i = 1, \ldots, k\} - \sum \eta(\Delta t_{\nu-1},\ n)\Delta t_{\nu-1} \\ \qquad \leqslant \Pr.\{y^{(n)}(s_\nu^{(i)}) \leqslant f(s_\nu^{(i)}),\ \nu = 1, \ldots, m,\ i = 1, \ldots, k\} \\ \qquad \leqslant \Pr.\{y^{(n)}(t_i) \leqslant f(t_i),\ i = 1, \ldots, k\}. \end{cases}$$

Now letting $\operatorname*{Max}_{i,\nu}(s_\nu^{(i)} - s_{\nu-1}^{(i)}) \to 0$, but $\rho(\Delta)$ being fixed, we see that the second member of (4.2) can be replaced by $\Pr.\{y^{(n)}(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1\}$. Next, in the inequality thus obtained, if we make $n \to \infty$ and use Theorem 2 we are led to

$$\Pr.\{y(t_i) \leqslant f(t_i) - 2\varepsilon,\ i = 1, \ldots, k\} - \eta[\rho(\Delta)]$$

$$\leqslant \varliminf_{n\to\infty} \Pr.\{y^{(n)}(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1\}$$

$$\leqslant \varlimsup_{n\to\infty} \Pr.\{y^{(n)}(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1\}$$

$$\leqslant \Pr.\{y(t_i) \leqslant f(t_i),\ i = 1, 2, \ldots, k\},\ \eta[\rho(\Delta)] \to 0, \text{ as } \rho(\Delta) \to 0.$$

Finally, taking into consideration of continuity of $y(t)$, if we let $\rho(\Delta) \to 0$, and then $\varepsilon \to 0$, we see that

$$\lim_{n\to\infty} \Pr.\{y^{(n)}(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1\} = \Pr.\{y(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1\}.$$

This proves the required result.

REMARK. In the above we have assumed continuity of the distribution of $y(t_i)$ and that $\Pr.(y(\tau) < f(\tau),\ 0 \leqslant \tau \leqslant 1) = \Pr.(y(\tau) \leqslant f(\tau),\ 0 \leqslant \tau \leqslant 1)$. If these were not true certain changes would be necessary. The former is true for many cases of pratical importance, and the latter was rigorously proved by R. Fortet [8] for the case $b = 1$.

Theorem 3 is an important case of the invariance principle, other cases will be proved using this theorem or directly in easier manner as in Donsker [2], Kac [12]. (c.f. also Udagawa [16]).

Next we shall apply Theorem 3 to the proof of Doob's argument stated in the beginning of this section. We shall begin by proving

THEOREM 4. *Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be order statistics from a population with continuous population distribution $F(x)$. Then $X_{(i)}$ becomes a Markov sequence, and*

$$n(t) = \sum_{i=1}^{n} \Phi_t(X_{(i)}), \quad -\infty < t < \infty$$

is a Markov process, where

$$\begin{aligned} \Phi_t(x) &= 1 \quad \text{for} \quad x \leqq t \\ &= 0 \quad \text{for} \quad 0 > t. \end{aligned}$$

Proof. Suppose, for simplicity, that the population distribution has the continuous density function $F'(x) = f(x)$, then

$$\Pr.\{x_k \leqq X_{(k)} \leqq x_k + dx_k \mid X_{(1)} = x_1, X_{(2)} = x_2, \ldots, X_{(k-1)} = x_{k-1}\}$$
$$= (n-k+1) f(x_k)\, dx_k \frac{\left(\int_{x_k}^{\infty} f(x)\,dx\right)^{n-k}}{\left(\int_{x_{k-1}}^{\infty} f(x)\,dx\right)^{n-k+1}}, \quad x_{k-1} \leqq x_k.$$

This means that $X_{(i)}$ is a Markov sequence with the transition probability

$$\Pr.\{X_{(k)} \leqq x_k \mid X_{(k-1)} = x_{k-1}\} = 1 - \frac{\left(1 - \int_{-\infty}^{x_k} f(x)\,dx\right)^{n-k+1}}{\left(1 - \int_{-\infty}^{x_{k-1}} f(x)\,dx\right)^{n-k+1}},$$

Also, for $0 < s < t$, $0 \leqq p \leqq q \leqq n$,

$$(4.3) \quad \left\{ \begin{aligned} &\Pr.\{n(t) = q \mid n(\tau) = g(\tau),\ \tau \leqq s\} \\ &= \binom{n-p}{q-p} \frac{\left(\int_s^t f(u)\,du\right)^{q-p} \left(\int_t^{\infty} f(u)\,du\right)^{n-q}}{\left(\int_s^{\infty} f(u)\,du\right)^{n-p}} \\ &= \Pr.\{n(t) = q \mid n(s) = g(s)\}, \end{aligned} \right.$$

where $g(\tau)$ is a step function with p jumps up to $\tau = s$, each jump being of magnitude 1. This relation shows that $n(t)$ is a Markov process with the transitions probability indicated by (4.3).

THEOREM 5. *With the same notations as in Theorem 4 we have*

$$(4.4) \qquad \Pr.\{\underset{-\infty<t<\infty}{\text{l. u. b.}} |F_n(t) - F(t)|\sqrt{n} \leqslant \alpha\} \to \Pr.\{\underset{0\leq t\leq 1}{\text{Max}} |y(\tau)| \leqslant \alpha\}, \quad n \to \infty,$$

where $F_n(t) = n(t)/n$, $y(\tau)$ *is a Gaussian process such that* $y(0) = y(1) = 0$, $E(y(u)y(v)) = \text{Min}(u, v) - uv$, *or it can be written as* $y(u) = (1-u)x\left(\frac{u}{1-u}\right)$, $0 \leqslant u \leqslant 1$, *by the Wiener process* $x(u)$.

Proof. Since, as is well known, (4.4) is free from the population distribution we may assume, simply, that it to be a uniform distribution over (0, 1). Then by (4.3),

$$(4.5) \quad \Pr.\{n(t) = q \,|\, n(s) = p\} = \binom{n-p}{q-p} \rho^{q-p}(1-\rho)^{n-q}; \ \rho = \frac{\Delta s}{1-s}, \ \Delta s = t - s.$$

If we write

$$(4.6) \qquad y^{(n)}(t) = (n(t) - n\cdot t)/\sqrt{n}$$

we are sufficed to show that $y^{(n)}(t)$ satisfies $(C_1') - (C_3')$ of Theorem 2.

First, by (4.5), (4.6)

$$(4.7) \quad \left\{ \begin{aligned} &\Pr.\{|y^{(n)}(t) - x| > \delta(1+x^2)^{1/2} \,|\, y^{(n)}(s) = x\} \\ &= \Bigg(\sum_{n(t-s)+\sqrt{n}\delta(1+x^2)^{1/2} \leq \alpha \leq n-p} + \sum_{0\leq \alpha \leq n(t-s)-\sqrt{n}\delta(1+x^2)^{1/2}} \Bigg) \\ &\quad \cdot \binom{n-p}{\alpha} \rho^{\alpha}(1-\rho)^{n-p-\alpha}, \quad p = ns + \sqrt{n}x. \end{aligned} \right.$$

Since x must satisfy the relations $n(t-s) + \sqrt{n}\delta(1+x^2)^{1/2} \leqslant n - p$, $n(t-s) - \sqrt{n}\delta(1+x^2)^{1/2} \geqslant 0$, we get in either case $n - p > nc(\mu)$, when $\rho(\Delta)$ is sufficiently small and $0 < t \leqslant 1 - \mu$, $\mu > 0$ being arbitrary but fixed. Therefore $n - p \to \infty$ uniformly in x, if $\mu > 0$ is fixed. Introduce the normalized variable $\xi = (\alpha - (n-p)\rho)/\sqrt{(n-p)\rho(1-\rho)}$, then the lower bound ξ_1 and upper bound ξ_2, which correspond to those of α in the first and second summation of (4.7) satisfy, when $\rho(\Delta)$ is small and $0 < t \leqslant 1 - \mu$,

$$\xi_1 > c_1(\mu)\frac{\delta(1+x^2)^{1/2}}{\Delta s}, \qquad \xi_2 < -c_2(\mu)\frac{\delta(1+x^2)^{1/2}}{\Delta s}.$$

Hence

$$\Pr.\{|y^{(n)}(t)-x|>\delta(1+x^2)^{1/2}\,|\,y^{(n)}(s)=x\}$$

$$\leq \frac{2}{\sqrt{2\pi}}\int_{c(\mu)\delta(1+x^2)^{1/2}/\Delta s}^{\infty} e^{-u^2/2}du \leq c'(\mu)\frac{(\Delta s)^2}{\delta^2(1+x^2)}.$$

which proves (C_1'), for the interval $(0,\ 1-\mu)$ instead of $(0,\ 1)$.

Next, by (4.5), it is easy to show that

$$E\{y^{(n)}(t)-x\,|\,y^{(n)}(s)=x\} = -x\frac{\Delta s}{1-s}$$

and

$$E\{(y^{(n)}(t)-x)^2\,|\,y^{(n)}(s)=x\}$$

$$=\Delta s-\frac{x}{\sqrt{n}}\frac{\Delta s}{1-s}+x^2\frac{(\Delta s)^2}{(1-x)^2}-\frac{(\Delta s)^2}{1-s}+\frac{x}{\sqrt{n}}\frac{(\Delta s)^2}{(1-s)^2}$$

$$=\Delta s+\varepsilon(\Delta s,\ n,\ x)\Delta s(1+x^2)^{1/2},$$

where $\varepsilon(\Delta s, n, x)$ satisfies the conditions in Theorem 2, in the interval $0<s<1-\mu$. Thus we have proved the « non-truncated form », in that restricted interval, i.e. the one with $\int_{|y-x|<\delta(1+x^2)^{1/2}} dF^{(n)}$ replaced by $\int_{-\infty}^{\infty} dF^{(n)}$, on the left-hand sides of (C_2'), (C_3'). If the non-truncated forms of (C_2'), (C_3') are satisfied, Theorem 2, 3 hold as well as with the truncated ones, as will be seen from the arguments in the proofs of these theorems. It would not, however, be difficult to deduce the truncated form of (C_2'), (C_3') directly from (4.5). The corresponding functions a, b now take the form $a(s,\ x)=-x/(1-s)$, $b(s,\ x)=1$. We have proved thus, according to Theorem 3,

$$\text{(4.8)}\quad \left\{\begin{array}{l}\Pr.\{\underset{0\leq\tau\leq 1-\mu}{\text{l. u. b.}}\ |y^{(n)}(\tau)|\leq\alpha\} \\ \to \Pr.\{\underset{0\leq\tau\leq 1-\mu}{\text{l. u. b.}}\ |y(\tau)|\leq\alpha\},\quad 0<\mu<1.\end{array}\right.$$

where $y(t)$ is the solution of (1.1) with a, b indicated in the above, or it is determined by the equivalent symbolic equation

$$dy(t)=-\frac{y(t)}{1-t}dt+dx(t),$$

which is easily seen to be satisfied by $y(t)$ in the statement of Theorem 5.

To extend (4.8) to the case $\mu = 0$, first note that $y^{(n)}(t)$ and $y(t)$ are symmetric about $t = 0$, and observe that

$$\Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\} - [1 - \Pr.\{\underset{0\le t\le \varepsilon}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\}]$$

$$= \Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\} - \Pr.\{\underset{1-\varepsilon\le t\le 1}{\text{l. u. b.}}\ |y^{(n)}(t)| > \alpha\}$$

$$\leqslant \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\} \leqslant \Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\}.$$

Hence, on making $n \to \infty$, by (4.8)

$$(4.9)\quad \begin{cases} \Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} - [1 - \Pr.\{\underset{0\le t\le \varepsilon}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\}] \\ \qquad \leqslant \overline{\lim_{n\to\infty}} \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\} \\ \qquad \leqslant \Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\}. \end{cases}$$

Also, given η, if $\varepsilon > 0$ is small

$$(4.10)\qquad \Pr.\{\underset{0\le t\le \varepsilon}{\text{l. u. b.}}\ |y(t)| > \alpha\} < \eta,$$

and so

$$(4.11)\quad \begin{cases} \Pr.\{\underset{0\le t\le 1-\varepsilon}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} \leqslant \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} \\ + \Pr.\{\underset{0\le t\le \varepsilon}{\text{l. u. b.}}\ |y(t)| > \alpha\} \leqslant \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} + \eta. \end{cases}$$

Substituting (4.10), (4.11) into (4.9) we get

$$\Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} - \eta$$

$$\leqslant \overline{\lim_{n\to\infty}} \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y^{(n)}(t)| \leqslant \alpha\}$$

$$\leqslant \Pr.\{\underset{0\le t\le 1}{\text{l. u. b.}}\ |y(t)| \leqslant \alpha\} + \eta,$$

which proves the theorem.

5. *Structure and uniqueness of the continuous Markov process.* - In this section we shall show that the Markov process satisfying (C_1) — (C_3) is uniquely determined and indeed it is the unique solution of Itō's equation (1.1). To treat this problem rigorously, we are needed however, to formulate (C_1) — (C_3) in a more precise manner.

Suggested by Theorem 3 and Itō's theorem we state the continuity conditions:

$$(C_1^{\cdot})\quad \int_{|y-x|>\delta(1+x^2)^{1/2}} dF(s,\ x:\ s+\Delta s,\ y)=\eta(\Delta s,\ x)\Delta s,$$

$$(C_2^{\cdot})\quad \int_{|y-x|\leq\delta(1+x^2)^{1/2}} (y-x)\,dF(s,\ x;\ s+\Delta s,\ y)=a(s,\ x)\Delta s+\varepsilon(\Delta s,\ x)(1+x^2)^{1/2}\Delta s,$$

$$(C_3^{\cdot})\quad \int_{|y-x|\leq\delta(1+x^2)^{1/2}} (y-x)^2\,dF(s,\ x;\ s+\Delta s,\ y)=b^2(s,\ x)\Delta s+\varepsilon(\Delta s,\ x)(1+x^2)\Delta s,$$

where $\eta(\Delta s,\ x)$, $\varepsilon(\Delta s,\ x)$, $a(t,\ x)$, $b(t,\ x)$ satisfy the conditions: given $\delta>0$, as $\Delta s\to 0$ (i) $\eta(\Delta s,\ x)\to 0$, uniformly in s, x, $0\leqslant s\leqslant 1$, $-\infty<x<\infty$, (ii) $\varepsilon(\Delta s,\ x)\to 0$, uniformly in s, $0\leqslant s\leqslant 1$, and in x belonging to every finite interval, (iii) $\varepsilon(\Delta s,\ x)$, $\eta(\Delta s,\ x)$ are bounded functions, (iv) $a(t,\ x)$, $b(t,\ x)$ are continuous in t, x, $0\leqslant t\leqslant 1$, $-\infty<x<\infty$, and satisfy $|a|+|b|\leqslant c(1+x^2)^{1/2}$.

When $E(y^2(t))<\infty$, we might also state the conditions in non-truncated forms;

$$(C_2^{\cdot\cdot})\quad \int_{-\infty}^{\infty} (y-x)\,dF(s,\ x;\ s+\Delta s,\ y)=a(s,\ x)\Delta s+\varepsilon(\Delta s,\ x)(1+x^2)^{1/2}\Delta s$$

$$(C_3^{\cdot\cdot})\quad \int_{-\infty}^{\infty} (y-x)^2\,dF(s,\ x;\ s+\Delta s,\ y)=b^2(s,\ x)\Delta s+\varepsilon(\Delta s,\ x)(1+x^2)\Delta s.$$

However, under the uniformity indicated by (i)-(iv), we can show that these two sets of conditions, $(C_1^{\cdot})$, $(C_2^{\cdot})$, $(C_3^{\cdot})$ and $(C_1^{\cdot})$, $(C_2^{\cdot\cdot})$, $(C_3^{\cdot\cdot})$ are equivalent. We can thus state

THEOREM 6. *Let* η, ε, a, b *satisfy* (i)-(iv) *in the above. If the Markov process* $y(t)$ *satisfies* $(C_1^{\cdot})$, $(C_2^{\cdot})$, $(C_3^{\cdot})$, then it satisfies $(C_1^{\cdot})$, $(C_2^{\cdot\cdot})$, $(C_3^{\cdot\cdot})$, *and conversely.*

Proof. Suppose $y(t)$ satisfies $(C_1^{\cdot})$-$(C_3^{\cdot})$. To prove the first half of the lemma introduce a division $\Delta(t_0,\ \ldots,\ t_n)$ of $(s,\ t)$ and a function

$$\Phi_\lambda(x)=\lambda\arc\tan\frac{x}{\lambda}$$

where $\lambda>0$ is a parameter tending to ∞, and write

$$(5.1)\qquad \begin{cases}\Phi_\lambda(y_\nu)=\Omega_\lambda(y_\nu)-\Phi_\lambda(y_{\nu-1})+\Phi_\lambda(y_{\nu-1})=\Delta\Phi+\Phi\\ \Phi_\lambda^4(y_\nu)=\Phi^4+4\Phi^3\Delta\Phi+6\Phi^2(\Delta\Phi)^2+4\Phi(\Delta\Phi)^3+(\Delta\Phi)^4.\end{cases}$$

Since for sufficiently small $\delta>0$, large $|x|$, and $|y-x|\leqslant\rho(\delta,\ x)=\delta(1+x^2)^{1/2}$

$$\Phi_\lambda(y)-\Phi_\lambda(x)=O\Big(\frac{|y-x|}{1+x^2/\lambda^2}\Big)=(y-x)\frac{1}{1+x^2/\lambda^2}+O\Big(\frac{(y-x)^2}{(1+x^2/\lambda^2)^2}\frac{|x|}{\lambda^2}\Big),$$

we get by (C_1^*)-(C_3^*)

$$
(5.2)\quad \left\{
\begin{aligned}
E\{\Delta\Phi(y_{\nu-1})|y_{\nu-1}\} &= \left(\int_{|y_\nu - y_{\nu-1}|\le\varrho(\delta, y_{\nu-1})} + \int_{|y_\nu - y_{\nu-1}|>\varrho(\delta, y_{\nu-1})}\right) \\
&\quad \cdot (\Phi_\lambda(y) - \Phi_\lambda(y_{\nu-1}))\,dF(t_{\nu-1}, y_{\nu-1}; t_\nu, y) \\
&= c_1\lambda\eta\,\Delta t_{\nu-1} + \int_{|y_\nu - y_{\nu-1}|\le\varrho(\delta, y_{\nu-1})} \frac{y - y_{\nu-1}}{1+\dfrac{y_{\nu-1}^2}{\lambda^2}}\,dF \\
&\quad + c_2\int_{|y - y_{\nu-1}|\le\varrho(\delta, y_{\nu-1})} \left(\frac{y-y_{\nu-1}}{1+\dfrac{y_{\nu-1}^2}{\lambda^2}}\right)^2 \frac{|y_{\nu-1}|}{\lambda^4}\,dF \\
&= c_1\lambda\,\Delta t_{\nu-1} + c_3\frac{(1+y_{\nu-1}^2)^{1/2}}{1+\dfrac{y_{\nu-1}^2}{\lambda^2}}\,\Delta t_{\nu-1} \\
&= \begin{cases} c_1\lambda\eta\,\Delta t_{\nu-1} + O(1+y_{\nu-1}^2)^{1/2}\Delta t_{\nu-1}, \\ c_1\lambda\eta\,\Delta t_{\nu-1} + O(1+\Phi(y_{\nu-1}))\Delta t_{\nu-1}, \end{cases} \\
&\qquad \eta\to 0 \quad \text{as} \quad \rho(\Delta)\to 0,
\end{aligned}\right.
$$

and similarly for $2 \leqslant k \leqslant 4$

$$
(5.3)\qquad E\{|\Delta\Phi|^k|y_{\nu-1}\} \leqslant \begin{cases} c\lambda^k\eta\,\Delta t_{\nu-1} + O(1+y_{\nu-1}^2)^{k/2}\Delta t_{\nu-1}, \\ c\lambda^k\eta\,\Delta t_{\nu-1} + O(1+\Phi_\lambda^2(y_{\nu-1}))^{k/2}\Delta t_{\nu-1}. \end{cases}
$$

Using the second estimate in (5.2) and (5.3), we get

$$|E(\Phi^3(y_\nu)\Delta\Phi(y_{\nu-1}))| \leqslant c\lambda^4\eta\,\Delta t_{\nu-1} + c(1+\mu_{\nu-1})\Delta t_{\nu-1}, \qquad \mu_\nu = E\Phi_\lambda^4(y_\nu),$$

and the same relation for the remaining last three terms of (5.1). Hence

$$\mu_\nu \leqslant \mu_{\nu-1} + c\lambda^4\eta\,\Delta t_{\nu-1} + c(1+\mu_{\nu-1})\Delta t_{\nu-1},$$

where $\eta\to 0$, as $\rho(\Delta)\to 0$. Successive substitution then gives

$$\mu_n \leqslant y_0^4\prod_{\nu=1}^n(1+c\,\Delta t_{\nu-1}) + c(1+\lambda^4\eta)\sum_{\nu=1}^n\Delta t_{\nu-1}.$$

Making $\rho(\Delta)\to 0$, since $\eta\to 0$ as $\rho(\Delta)\to 0$

$$E\{\Phi_\lambda^4(y(t))|y(s)=y_0\} \leqslant cy_0^4 + c.$$

Finally, letting $\lambda\to\infty$, we get

$$(5.4)\qquad E(y^4(t)|y(s)=y_0) \leqslant c'(1+y_0^2)^2.$$

Next write, with obvious notations

$$
(5.5)\qquad \begin{cases} S_n = \sum_{\nu=1}^{n} \{\Phi_\lambda(y_\nu) - \Phi_\lambda(y_{\nu-1})\}, \quad S_n = S_{n-1} + \Delta\Phi(y_{n-1}), \\ S_\nu^4 = S_{\nu-1}^4 + 4S_{\nu-1}^3 \Delta\Phi(y_{\nu-1}) + 6S_{\nu-1}^2(\Delta\Phi(y_{\nu-1}))^2 \\ \qquad + 4S_{\nu-1}(\Delta\Phi(y_{\nu-1}))^3 + (\Delta\Phi(y_{\nu-1}))^4. \end{cases}
$$

Then by the first estimates in (5.2), (5.3) and (5.4), we get

$$
\begin{aligned} |ES_{\nu-1}^3 \Delta\Phi(y_{\nu-1})| &\leqslant c\lambda\eta\Delta t_{\nu-1}(ES_{\nu-1}^4)^{3/4} + c(ES_{\nu-1}^4)^{3/4}(E(1+y_{\nu-1}^2)^2)^{1/4}\Delta t_{\nu-1} \\ &\leqslant c'\lambda\eta\Delta t_{\nu-1}\mu_{\nu-1}'^{3/4} + c'\mu_{\nu-1}'^{3/4}(1+y_0^2)^{1/2}, \quad \mu_\nu' = ES_\nu^4. \end{aligned}
$$

Similarly for the remaining three terms of (5.5)

$$
|ES_{\nu-1}^{4-k}[\Delta\Phi(y_{\nu-1})]^k| \leqslant c'\lambda^k\eta\Delta t_{\nu-1}\mu_{\nu-1}'^{(4-k)/4} + c'\mu_{\nu-1}'^{(4-k)/4}(1+y_0^2)^{k/2}\Delta t_{\nu-1},
$$
$$
2 \leqslant k \leqslant 4.
$$

Substitute these into (5.5), then we get

$$
\mu_\nu' \leqslant \mu_{\nu-1}'(1 + c'\lambda^4\eta\Delta t_{\nu-1}) + \mu_{\nu-1}'^{3/4}c(y_0)\Delta t_{\nu-1} + \gamma\Delta t_{\nu-1},
$$
$$
c(y_0) = c(1+y_0^2)^2, \qquad \gamma = c'\lambda^4\eta + c(y_0).
$$

Obviously μ_ν' are majorated by the solutions of the difference equation

$$
u_\nu = u_{\nu-1}(1 + c'\lambda^4\eta\Delta t_{\nu-1}) + u_{\nu-1}^{3/4}c(y_0)\Delta t_{\nu-1} + \gamma\Delta t_{\nu-1}.
$$

If we let $\rho(\Delta) \to 0$, then, since $\eta \to 0$, we are led to the differential equation

$$
\frac{du}{d\tau} = (1 + u^{3/4})c(y_0), \qquad s \leqslant \tau \leqslant t,
$$

of which $E\{\Phi_\lambda(y(\tau)) - \Phi_\lambda(y_0)\}^4$ is majorated by the solution satisfying $u(s) = 0$. We have thus, at least for small value of $|t-s|$,

$$
E\{\Phi_\lambda(y(t)) - \Phi_\lambda(y_0)\}^4 \leqslant c(1+y_0^2)^2(t-s),
$$

and so

$$
(5.6)\qquad E(y(t) - y_0)^4 \leqslant c(1+y_0^2)^2(t-s), \qquad y(s) = y_0.
$$

Thus prepared, it is immediate to deduce (C_2^{**}), (C_3^{**}). First

$$
\begin{aligned} &\int_{-\infty}^{\infty} (y-x)^2 dF(s, x; s+\Delta s, y) \\ &= \left(\int_{|y-x|\leq\rho(\delta,x)} + \int_{|y-x|>\rho(\delta,x)}\right)(y-x)^2 dF \\ &= b^2(s, x)\Delta s + \varepsilon(\Delta s, x)(1+x^2)\Delta s \\ &\quad + O\left\{\left(\int_{-\infty}^{\infty}(y-x)^4 dF\right)^{1/2}\left(\int_{|y-x|>\rho(\delta,x)} dF\right)^{1/2}\right\}. \end{aligned}
$$

But this is, in view of (5.6), smaller than

$$b^2(s, x)\Delta s + \varepsilon(\Delta s, x)(1+x^2)\Delta s + (1+x^2)\eta(\Delta s, x)^{1/2}\Delta s,$$

which is obviously just the form on the right-hand side of (C_2^{**}). Next (C_3^{**}) is derived from (C_2^{**}) in a similar manner. This completes the proof of the first half of the theorem.

To prove the second half, suppose (C_1^{*}), (C_2^{**}), (C_3^{**}) are satisfied. Then we have

$$(5.7)\quad \begin{cases} \displaystyle\int_{|y-x|\leq\varrho(\delta,x)} (y-x)^2 dF(s, x;\, s+\Delta s, y) \\ \displaystyle\qquad \leqslant \int_{-\infty}^{\infty} (y-x)^2 dF \leqslant c(1+x^2)\Delta s, \end{cases}$$

and so

$$(5.8)\quad \begin{cases} \displaystyle\int_{|y-x|\leq\varrho(\delta,x)} (y-x)\, dF(s, x;\, s+\Delta s, y) \\ \displaystyle = \left(\int_{-\infty}^{\infty} - \int_{|y-x|>\varrho(\delta,x)}\right)(y-x)\, dF \\ = a\Delta s + \varepsilon(\Delta s, x)(1+x^2)^{1/2}\Delta s \\ \displaystyle\qquad + O\left\{\left(\int_{-\infty}^{\infty}(y-x)^2 dF\right)^{1/2}\left(\int_{|y-x|\geq\varrho(\delta,x)} dF\right)^{1/2}\right\} \\ = a\Delta s + \varepsilon(\Delta s, x)(1+x^2)^{1/2}\Delta s + c'(1+x^2)^{1/2}\eta(\Delta s, x)^{1/2}\Delta s, \end{cases}$$

which is the same form as the right-hand side of (C_2^{*}). We have thus (C_2^{*}) and (5.7) which is weaker than (C_3^{*}). From (C_1^{*}), (C_2^{*}), (5.7) we can deduce, however, the relation (5.6). For when we deduced (5.6) in the above, we have only used this weaker relation (5.7), instead of (C_3^{*}) itself. Then representing the integral $\int_{|y-x|\leq\varrho(\delta,x)} (y-x)^2 dF$ as a difference of two integrals, as in (5.8), and using (5.6) we get (C_3^{*}). We have thus proved the theorem, and are in a position to state the following theorem on the uniqueness and representation of the Markov process. A similar theorem has recently been published by J. L. Doob [4]. His formulation is somewhat different from this, and assumptions on a, b are weaker than ours. When I independently obtained this theorem in my preliminary report [15], the uniform Lipschitz conditions were imposed on a, b. But afterwards, suggested by Doob's result, I tried to remove them and found that my original proof applies without essential changes to the generalized case. This depends on the fact that in this theorem the existence of the Markov process $y(t)$ is presupposed, whereas in Theorem 1 the Lipschitz conditions are needed to prove it. Doob's approach depends on the martingale. He treats the case $b=0$, which escaped from our analysis.

According to our theorem and Itō's theorem, when a, b moreover satisfy the Lipschitz condition, (C_1^{*})-(C_3^{*}) determine $y(t)$ uniquely.

THEOREM 7. *Suppose that* $a(t, x)$, $b(t, x)$, $\varepsilon(\Delta s, x)$, $\eta(\Delta s, x)$ *satisfy the conditions under Theorem* 2. *Then a necessary condition that* $y(t)$ *should satisfy Itō's equation* (1.2) *is that the transition probability should satisfy* (C_1^{*})-(C_3^{*}) *or equivalently* $(C_1^{'})$, $(C_2^{''})$, $(C_3^{''})$. *Conversely these conditions are sufficient for* $y(t)$ *to be a solution of* (1.2), *provided that* $b(t, x)>0$.

Almost all sample functions of $y(t)$ *are then continuous.*

We require the following lemmas.

LEMMA 10. *Let* $a(t, x)$, $b(t, x)$ *be continuous and bounded* $|a|$, $|b|<M$, *and* $y(t)$ *be Markov process satisfying the equation* (1.2). *Then, for every real* s

$$E(e^{sy(t)}) \leqslant c(s)\, e^{sy_0}, \tag{5.9}$$

$$E \exp\left[s\int_0^t f(\tau, y)\, dx(\tau) - \frac{s^2}{2}\int_0^t f^2(\tau, y)\, d\tau\right] = 1, \tag{5.10}$$

where $f(t, x)$ *is a bounded continuous function.*

Proof. Put

$$P_n = \prod_{\nu=0}^{n} [1 + s\, a(t_\nu, y_\nu)\Delta t_\nu + s\, b(t_\nu, y_\nu)\psi(\Delta x_\nu)],$$

where $y_\nu = y(t_\nu)$, $\Delta x_{\nu-1} = x(t_\nu) - x(t_{\nu-1})$, and $\psi(x)$ is the function introduced in Lemma 6. Then as in Lemma 6, we have for $p>0$

$$E(P_m^{2p}) \leqslant c_{2p}.$$

But, by the equation (1.2) we get

$$E(\Delta y(t))^2 \leqslant 2E\left(\int_t^{t+\Delta t} a(\tau, y)\, d\tau\right)^2 + 2E\left(\int_t^{t+\Delta t} b(\tau, y)\, dx\right)^2 \leqslant 4M^2\Delta t.$$

Hence in the same way as in the proof of Lemma 6 we see that there exists, in the L^2-sense, $P = \underset{\rho(\Delta)\to 0}{\mathrm{l.i.m.}}\, P_m$ and

$$E(P) = E \exp\left[s\int_0^t a\, d\tau - \frac{s^2}{2}\int_0^t b^2\, d\tau + s\int_0^t b\, dx(\tau)\right] \leqslant c_1,$$

whence

$$E \exp\left[s\int_0^t a\,dt + s\int_0^t b\,dx(\tau)\right] \leqslant c_1 e^{\frac{s^2}{2}M^2} = c(s). \tag{5.11}$$

This proves (5.9).

The second part is proved in the same way as with the first one, if we only note that for every division Δ

$$E\left\{\prod_0^n [1 + sf(t_\nu, y_\nu)\psi(\Delta x_\nu)]\right\} = 1$$

whose limit is nothing but the relation (5.10).

LEMMA 11. *Let $a(t, x)$, $b(t, x)$ be continuous functions satisfying*

$$|a(t, x)| + |b(t, x)| \leqslant c(1 + x^2)^{1/2},$$

while $\alpha(t, z)$, $\beta(t, z)$ be defined by

$$\alpha(t, z) = \frac{a(t, \sinh z)}{\cosh z} - \frac{\sinh z}{2\cosh^3 z} b^2(t, \sinh z), \tag{5.12}$$

$$\beta(t, z) = \frac{b(t, \sinh z)}{\cosh z}. \tag{5.13}$$

If $y(t)$ satisfies (1.2), then $z(t) = z(y(t))$, where $z(y)$ is the function introduced in Lemma 2, satisfies (1.2) with a, b replaced by α, β and vice versa. It holds that

$$E(y(t))^{2p} \leqslant c_p(1 + y_0^2)^p, \qquad E(e^{sz(t)}) \leqslant c(s)e^{sz_0} \tag{5.14}$$

for every $p > 0$, and real s.

This is an example of transformations of a Markov process into others. More general transformations containing the time parameter, e.g. $z(t) = \Phi[t, y(t)]$ may be considered. These transformations correspond to change of variables in the functional equation satisfied by the transition probability.

Proof. After the method in Itō [10], p. 35, define $a_N(t, y)$, $b_N(t, y)$ as follows:

$$\begin{aligned} a_N(t, y) &= a(t, y) && \text{for } |a(t, y)| \leqslant N, \\ &= N && \text{for } a(t, y) \geqslant N, \\ &= -N && \text{for } a(t, y) \leqslant -N, \end{aligned}$$

similarly for $b_N(t, y)$. Put now

$$y_N(t) = y_0 + \int_0^t a_N(t, y)\,d\tau + \int_0^t b_N(\tau, y)\,dx(\tau).$$

Then

$$
(5.15)\quad \left\{
\begin{aligned}
z[y_N(t)] &= z[y_N(0)] + \sum_{\nu=1}^{n}\{z[y_N(t_\nu)] - z[y_N(t_{\nu-1})]\} \\
&= z[y_N(0)] + \sum_{1}^{n}\frac{1}{\sqrt{1+y_N^2(t_{\nu-1})}}\left\{\int_{t_{\nu-1}}^{t_\nu} a_N(\tau, x)\,d\tau \right.\\
&\left. + \int_{t_{\nu-1}}^{t_\nu} b_N(\tau, y)\,dx\right\} - \sum_{1}^{n}\frac{y_N(t_{\nu-1})}{2[1+y_N^2(t_{\nu-1})]^{3/2}} \\
&\cdot\left\{\int_{t_{\nu-1}}^{t_\nu} a_N(\tau, y)\,d\tau + \int_{t_{\nu-1}}^{t_\nu} b_N(\tau, y)\,dx\right\}^2 + \sum_{1}^{n} O(|\Delta y_N|^3).
\end{aligned}
\right.
$$

Making $\rho(\Delta)\to 0$, we get

$$
(5.16)\quad \left\{
\begin{aligned}
z[y_N(t)] &= z[y_N(0)] + \int_0^t \left\{\frac{a_N(\tau, y)}{[1+y_N^2(\tau)]^{1/2}} - \frac{y_N(\tau)\,b_N^2(\tau, y)}{2[1+y_N^2(\tau)]^{3/2}}\right\} d\tau \\
&+ \int_0^t \frac{b_N(\tau, y)}{\sqrt{1+y_N^2(\tau)}}\,dx(\tau),
\end{aligned}
\right.
$$

the right-hand member of (5.15) tending to that in (5.16) in the L^2-sense. For, the first summation in the second term of (5.15), say S_1, for which holds $E(|S_1|^k)\leqslant N^k$, tends to the first integral in (5.16). Next we note that the second summation, say S_2, can be approximated by

$$
S_2' = \sum_{1}^{n}\int_{t_{\nu-1}}^{t_\nu}\frac{b_N(\tau, y)}{\sqrt{1+y_N^2(\tau)}}\,dx
$$

in such a way that we have

$$
E(S_2 - S_2')^2 = E\left\{\sum_{1}^{n}\left(\frac{1}{\sqrt{1+y_N^2(t_{\nu-1})}}\int_{t_{\nu-1}}^{t_\nu} b_N(\tau, y)\,dx - \int_{t_{\nu-1}}^{t_\nu}\frac{b_N(\tau, y)}{\sqrt{1+y_N^2(\tau)}}\,dx\right)\right\}^2,
$$

of which arising cross-products vanish, on taking expectation, and

$$
\begin{aligned}
&\leqslant N^2\sum\int_{t_{\nu-1}}^{t_\nu} E[y_N(t_{\nu-1}) - y_N(t)]^2\,dt \\
&\leqslant 4N^4\sum(\Delta t_{\nu-1})^2 \leqslant 4N^4\rho(\Delta)\to 0, \quad \rho(\Delta)\to 0.
\end{aligned}
$$

When we consider the third term in (5.15), we make use of

$$
(5.17)\quad \left\{
\begin{aligned}
&E\left(\int_u^v b_N(\tau, y)\,dx\right)^4 = 3E\left(\int_u^v b_N^2\,d\tau\right)^2 \\
&+ 6E\left\{\left(\int_u^v b_N\,dx\right)^2\left(\int_u^v b_N^2\,d\tau\right)\right\} \leqslant c(v-u)^2,
\end{aligned}
\right.
$$

a direct consequence of comparisons of powers in s of (5.10). Then if we expand the squared bracket, the terms containing only $(\int b_N dx)^2$ contribute to our computation, and they are approximated by $\int_{t_{\nu-1}}^{t_\nu} b_N^2 d\tau$ in such a way that

$$E\left\{\left(\int_{t_{\nu-1}}^{t} b_N(\tau, y)\,dx\right) - \int_{t_{\nu-1}}^{t_\nu} b_N^2(\tau, y)\,d\tau\right\}^2 \leqslant c(N)(\Delta t_{\nu-1})^2.$$

In the same way the last term is seen to vanish when $\rho(\Delta)\to 0$. Thus we have (5.16).

To proceed rigorously, we introduce the parameter ω, and write $y_N(t)=y_N(t, \omega)$, $\omega\in\Omega$, and remember that according to [9] there exists an ascending sequence of sets $\Omega_N \subset \Omega$, such that $P(\Omega_N)\uparrow 1$, and $y_N(t, \omega)=y(t, \omega)$ if $\omega\in\Omega_N$. Therefore (5.16) holds even if we take off the subscripts N, and then changing y to z in the equality thus obtained, $z(t)$ is now seen to satisfy (1.2) with a, b replaced by α, β.

That $z(t)$ satisfies the second inequality, whence $y(t)$ also the first, follows from Lemma 10.

To deduce conversely from the equation belonging to $z(t)$, that belonging to $y(t)$, we can proceed similarly, the approximation procedures being easier than those in the above.

Now we shall return to the proof of Theorem 7 and first prove

Necessity. Consider the transformed $z(t)$ together with $y(t)$ and put $\Delta z = z(s+\Delta s) - z(s)$, and similarly for $y(t)$. Then obtaining from (3.4), by repeated use of Schwarz's inequality,

$$\begin{aligned} E_x\{(\Delta y)^4\} &\equiv E\{(\Delta y)^4 \mid y(s) = x\} \\ &\leqslant \frac{1}{8}\{(E_x e^{8z(s)})^{1/2} + (E_x e^{-8z(s)})^{1/2}\}\{E_x(\Delta z)^{16}\}^{1/4}\{E_x e^{16|\Delta z|}\}^{1/4} \end{aligned}$$

and using

$$E_x(\Delta z)^{16} \leqslant c_1(\Delta s)^4, \qquad E_x e^{16|\Delta z|} \leqslant c_2,$$

which follow from (5.9) and (5.10) as in the proof of (5.17), we get

$$E_x(\Delta y)^4 \leqslant c_3(e^{4z_0} + e^{-4z_0})(\Delta t)^2 \leqslant c_4(1 + y_0^2)^2(\Delta s)^2.$$

So that the Tchebychev inequality gives

$$\text{Pr.}\{|\Delta y| > \delta(1 + x^2)^{1/2} \mid y(s) = x\} \leqslant \delta^{-4}(1 + x^2)^{-2} c_4(1 + x^2)^2(\Delta s)^2 = \delta^{-4} c_4(\Delta s)^2,$$

which proves that (C_1^*) holds.

Now

$$E\{\Delta y(s)\,|\,y(s)=x\}=E\left\{\int_s^{s+\Delta s}[a(\tau,\,y(\tau))-a(s,\,y(s))]\,d\tau\,|\,y(s)=x\right\}+a(s,\,x)\Delta s. \tag{5.18}$$

But, when $-\infty < x < \infty$, the first term can be written as

$$O\left\{\int_s^{s+\Delta s}[(1+x^2)^{1/2}+(E_x(1+y^2(\tau))^{1/2}]\,d\tau\right\}$$

which is by Lemma 11

$$\leqslant c(1+x^2)^{1/2}\Delta s. \tag{5.19}$$

Also, when $|x|\leqslant M$, the first term in (5.18) can be written as

$$O\left\{\int_s^{s+\Delta s}\left(\int_{|y-x|\leq\delta}+\int_{|y-x|>\delta}\right)|a(\tau,\,y)-a(s,\,x)|\,dF\cdot d\tau\right\}$$

$$\leqslant c(M)\varepsilon\Delta s+c(M)\int_s^{s+\Delta s}\left(\int_{-\infty}^{\infty}(1+y^2)\,dF\right)^{1/2}\left(\int_{|y-x|\geq\delta}dF\right)^{1/2}d\tau$$

$$=c(M)(1+x^2)\eta,$$

where ε, η is small if only Δs is small. This together with (5.19) proves (C_2^*).

(C_3^*) can be proved in the same way.

In the same way as in Lemma 2, when we use (3.4), (C_1^*), (C_2^{**}), (C_3^{**}) are transformed into the uniform cases. So that, by Lemma 11, for the remaining part of the proof, with no loss of generality, we may restrict ourselves to

Sufficiency in the uniform cases. Ist Step. We shall begin by constructing a Brownian motion.

Consider a division $\Delta(t_0,\,\ldots,\,t_n)$ of (0, 1). Let

$$\zeta_\nu=\frac{\Delta y_{\nu-1}-a_{\nu-1}\Delta t_{\nu-1}}{b_{\nu-1}},\quad y_\nu=y(t_\nu),\ a_\nu=a(t_\nu,\,y_\nu),\ b_\nu=b(t_\nu,\,y_\nu)$$

and let

$$S_\nu^*=\sum_{i=1}^{\nu}\zeta_i\lambda_{i-1}+\sum_{\nu+1}^{n}\xi_i,\quad 1\leqslant\nu\leqslant n-1,\quad \xi_i=\xi(t_i),$$

$$S_n^*\equiv\sum_1^n\xi_i,\quad S_0^*\equiv\sum_1^n\lambda_{i-1}\xi_i,\quad S_n=\sum_1^n\zeta_i,$$

$$\lambda_\nu=\lambda(y_\nu),$$

where $\xi(t)$ is a Brownian motion independent of $y(t)$, $\xi(0)=0$ $E(\Delta\xi(t))^2=\Delta t$, and

$$\lambda(x)=\begin{cases}1 & |x|\leq M,\\ 0 & |x|>M,\end{cases}$$

M a parameter tending to ∞. Then obviously

$$(5.20)\qquad \left|E\left\{\sum_1^n (e^{izS_\nu^*}-e^{izS_{\nu-1}^*})\right\}-E\{e^{izS_n}-e^{izS_0^*}\}\right|=|E(e^{izS_n^*}-e^{izS_n})|\leq\varepsilon(M),$$

where $\varepsilon(M)\to 0$, as $M\to\infty$.

Now observe that

$$(5.21)\qquad |E(e^{izS_\nu^*}-e^{izS_{\nu-1}^*})|\leq E\{|E\{(e^{iz\zeta_\nu\lambda_{\nu-1}}-e^{iz\xi_\nu})|y_{\nu-1}\}|\}$$

and write

$$(5.22)\qquad \begin{cases} E\{(e^{iz\zeta_\nu\lambda_{\nu-1}}-e^{iz\xi_\nu})|y_{\nu-1}\}\\ =\left(\int_{|y_\nu-y_{\nu-1}|>\delta}+\int_{|y_\nu-y_{\nu-1}|\leq\delta}\right)e^{iz\zeta_\nu\lambda_{\nu-1}}-e^{-\frac{z^2}{2}\Delta t_{\nu-1}}.\end{cases}$$

Then first by (C_1^*), the first integral is $\eta(\Delta t_{\nu-1})\Delta t_{\nu-1}$, and the second is, again by (C_1^*),

$$(5.23)\qquad \begin{cases} 1+\eta(\Delta t_{\nu-1})\Delta t_{\nu-1}+\int_{|y_\nu-y_{\nu-1}|\leq\varrho}[iz\zeta_\nu\lambda_{\nu-1}-\dfrac{z^2\zeta_\nu^2\lambda_{\nu-1}^2}{2}\\ \qquad +O(|\zeta_\nu^3\lambda_{\nu-1}^3|)]\,dF(t_{\nu-1},y_{\nu-1};t_\nu,y_\nu).\end{cases}$$

Second by (C_2^*) and (C_3^*)

$$(5.24)\qquad \int_{|y_\nu-y_{\nu-1}|\leq\delta}\zeta_\nu\lambda_{\nu-1}\,dF=\frac{\lambda_{\nu-1}}{b_{\nu-1}}\varepsilon(\Delta t_{\nu-1},y_{\nu-1})\Delta t_{\nu-1},$$

$$(5.25)\qquad \begin{cases}\int_{|y_\nu-y_{\nu-1}|\leq\delta}\zeta_\nu^2\lambda_{\nu-1}^2\,dF=\dfrac{\lambda_{\nu-1}^2}{b_{\nu-1}^2}\int_{|y_\nu-y_{\nu-1}|\leq\delta}(y_\nu-y_{\nu-1})^2\,dF\\ \qquad +c(M)((\Delta t_{\nu-1})^2+\delta\Delta t_{\nu-1})=\lambda_{\nu-1}\Delta t_{\nu-1}\\ \qquad +c(M)(\lambda_{\nu-1}\varepsilon(\Delta t_{\nu-1},y_{\nu-1})\Delta t_{\nu-1}+(\Delta t_{\nu-1})^2+\delta\Delta t_{\nu-1}),\end{cases}$$

where $c(x)$ is bounded in every finite x-interval, and also by (C_2^*)

$$(2.26)\qquad \left|\int_{|y_\nu-y_{\nu-1}|\leq\delta}\zeta_\nu^3\lambda_{\nu-1}^3\,dF\right|\leq c(M)\delta\Delta t_{\nu-1}.$$

If we substitute (5.23) together with (5.24), (5.25), (5.26), into (5.22) we get

$$(5.27)\quad \begin{cases} |E\{(e^{iz\zeta_\nu\lambda_{\nu-1}} - e^{iz\xi_\nu})|y_{\nu-1}\}| \leqslant (1-\lambda_{\nu-1})\Delta t_{\nu-1} \\ + c(M)[\lambda_{\nu-1}\varepsilon(\Delta t_{\nu-1}, y_{\nu-1})\Delta t_{\nu-1} + (\Delta t_{\nu-1})^2 + \delta\Delta t_{\nu-1}]. \end{cases}$$

Now, since the method of Lemma 9 yields

$$\text{Pr.}\{\max_{0\leq\nu\leq n}|y(t_\nu)| > k\} \to 0, \quad k \to \infty$$

uniformly in the choice of Δ, we have

$$(5.28)\qquad 0 \leqslant E(1-\lambda_{\nu-1}) \leqslant \varepsilon(M) \to 0, \quad M \to \infty.$$

Hence (5.21) combined with (5.27) yields

$$\overline{\lim_{\varrho(\Delta)\to 0}}\, |E(e^{izS_n}) - e^{-\frac{z^2}{2}}| \leqslant \lim_{M\to\infty}\lim_{\delta\to 0}\overline{\lim_{\varrho(\Delta)\to 0}}\{2\varepsilon(M)$$
$$+ c(M)E[\lambda_{\nu-1}\varepsilon(\Delta t_{\nu-1}, y_{\nu-1})] + \rho(\Delta) + \delta\} = 0.$$

Thus we have proved that S_n converges, in probability law, to a Gaussian random variable. In the same way we can show that if we form S_n from the $y(\tau)$, $t_0 \leqslant s \leqslant \tau \leqslant t \leqslant 1$, and if $\Psi(t_0)$ is an arbitrary functional of $y(\tau)$, $0 \leqslant \tau \leqslant t_0$, then

$$\text{Pr.}\{S_n \leqslant x | \Psi(t_0)\} \to \frac{1}{\sqrt{2\pi(t-s)}}\int_{-\infty}^{x} e^{-u^2/2(t-s)}du, \ \rho(\Delta) \to 0.$$

Therefore S_n over (s, t) determines a Brownian motion, $x(\tau)$, $s \leqslant \tau \leqslant t$, independent of $y(\tau)$, $0 \leqslant \tau \leqslant s$.

Consider a random function

$$f(t) = E\{\xi(t) - \xi(t_1)|y(t_0)\}, \quad t_0 \leqslant t_1 \leqslant t \leqslant 1,$$

where

$$\xi(t) = y(t) - \int_0^t a(\tau, y)\,d\tau.$$

Then, since a is bounded, we have by (C_2^{**})

$$|\Delta f(s)| = |E\{a(s, y(s))\Delta s + \varepsilon(\Delta s, y(s))\Delta s - \int_s^{s+\Delta s} a(\tau, y)\,d\tau|y(t_0)\}| \leqslant c\Delta s, \ s \geqslant t_1.$$

This means that $f(s)$ is absolutely continuous.

Now observe that for $h > 0$

$$E\left\{\frac{y(s+h) - y(s)}{h}\Big|y(t_0) = x\right\} = \int_{-\infty}^{\infty}\{a(s, y) + \varepsilon(h, y)\}\,dF(t_0, x; s, y)$$
$$\to \int_{-\infty}^{\infty} a(s, y)\,dF, \quad h \to 0,$$

by the bounded convergence of $\varepsilon(\Delta s, y)$ and also

$$\frac{1}{h}\int_s^{s+h} E\{a(\tau, y)\,|\,y(t_0) = x\}\,d\tau \to \int_{-\infty}^{\infty} a(s, y)\,dF, \quad h \to 0.$$

So that

$$h^{-1}\{f(s+h) - f(s)\} \to 0, \quad \text{as} \quad h \to +0.$$

Hence by absolute continuity of $f(t)$, $f(t) = \text{const.}$, $t_1 \leqslant t \leqslant 1$. But, since $f(t_1) = 0$, we have

(5.29) $$f(t) = 0, \qquad t_1 \leqslant t \leqslant 1.$$

2nd Step. In this step we represent $y(t)$ as a solution of Itō's integral equation, with the Brownian motion introduced in the first step.

Let $\Delta(t_0, t_1, \ldots, t_n)$ be a division of $(0, t)$, $\Delta_\nu(u_{\nu 0}, u_{\nu 1}, \ldots, u_{\nu m_\nu})$ sub-divisions of each sub-interval $(t_{\nu-1}, t_\nu)$,

$$u_{\nu 0} = t_{\nu-1} < u_{\nu 1} < \cdots < u_{\nu m_\nu} = t_\nu, \qquad \nu = 1, 2, \ldots, n,$$

and let us write

(5.30) $$\begin{cases} y(t) = y_0 + \sum_{\nu=1}^{n}\{y(t_\nu) - y(t_{\nu-1})\} \\ \quad = y_0 + \int_0^t a(\tau, y(\tau))\,d\tau \\ \qquad + \sum_{\nu=1}^{n} b_{\nu-1} S_{\Delta_\nu} + \sum_{\nu=1}^{n}\sum_{\mu=1}^{m_\nu} X_{\nu,\mu}, \end{cases}$$

where

$$\xi_{\nu,\mu} = \xi(u_{\nu,\mu}), \quad b_{\nu,\mu} = b[u_{\nu,\mu}, y(u_{\nu,\mu})], \quad b_\nu = b[t_\nu, y(t_\nu)],$$

$$X_{\nu,\mu} = \frac{\xi_{\nu,\mu} - \xi_{\nu,\mu-1}}{b_{\nu,\mu-1}}(b_{\nu,\mu-1} - b_{\nu-1}), \quad S_{\Delta_\nu} = \sum_{\mu=1}^{m_\nu}\frac{\xi_{\nu,\mu} - \xi_{\nu,\mu-1}}{b_{\nu,\mu-1}}.$$

First we shall prove convergence of the last term in (5.30), in probability, to zero. For this purpose we write it as

$$\sum_\nu\sum_\mu X_{\nu,\mu}\lambda_{\nu,\mu-1} + \sum_\nu\sum_\mu X_{\nu,\mu}(1 - \lambda_{\nu,\mu-1}),$$

where $\lambda_{\nu,\mu} = \lambda[y(u_{\nu,\mu})]$, and observe that

$$E\Big(\sum_\nu\sum_\mu X_{\nu,\mu}\lambda_{\nu,\mu-1}\Big)^2 = \sum_\nu\sum_\mu E(X^2_{\nu,\mu}\lambda^2_{\nu,\mu-1}),$$

since, on taking expectations, the arising cross products vanish by (5.29), and also by (C_3''')

$$E(X_{\nu,\mu}^2 \lambda_{\nu,\mu-1}^2) \leqq c_1(M) E[(b_{\nu,\mu-1} - b_{\nu-1})^2 E\{(\xi(u_{\nu,\mu}) - \xi(u_{\nu,\mu-1}))^2 | y(u_{\nu,\mu-1})\}]$$
$$\leqq c_2(M)(u_{\nu,\mu} - u_{\nu,\mu-1}) E\{b[u_{\nu,\mu-1}, y(u_{\nu,\mu-1})] - b[t_{\nu-1}(y(t_{\nu-1})]^2\}.$$

But, since $E[y(t)]^2 \leqq c$, $E[\Delta y(t)]^2 \leqq c\Delta t$ as shown in Theorem 7, we have, in general, $E\{b[s, y(s)] - b[s+\Delta s, y(s+\Delta s)]\}^2 \leqq \varepsilon(\Delta s)$, where $\varepsilon(\Delta s) \to 0$, as $\Delta s \to 0$ uniformly in other variables. Therefore

$$(5.31) \qquad \begin{cases} E\left(\sum_\nu \sum_\mu X_{\nu,\mu} \lambda_{\nu,\mu-1}\right)^2 \leqq c(M)\varepsilon[\rho(\Delta)] \sum_\nu \Delta t_{\nu-1} \\ \leqq c(M)\varepsilon[\rho(\Delta)], \quad \varepsilon[\rho(\Delta)] \to 0, \quad \text{as} \quad \rho(\Delta) \to 0. \end{cases}$$

On the other hand, by Lemma 9, since

$$\Pr.\left\{\sum_\nu \sum_\mu |1 - \lambda_{\nu,\mu-1}| \neq 0\right\} \leqq \varepsilon(M), \quad \varepsilon(M) \to 0, \quad \text{as} \quad M \to \infty,$$

we have, for every $\delta > 0$

$$(5.32) \qquad \Pr.\left\{\left|\sum_\nu \sum_\mu X_{\nu,\mu}(1 - \lambda_{\nu,\mu-1})\right| > \delta\right\} < \varepsilon(M).$$

This together with (5.31) gives

$$\Pr.\left\{\left|\sum_\nu \sum_\mu X_{\nu,\mu} \lambda_{\nu,\mu-1}\right| > 2\delta\right\} \leqq \delta^{-2} c(M)\varepsilon[\rho(\Delta)] + \varepsilon(M),$$

whence

$$(5.33) \qquad \operatorname*{st.\,lim}_{\rho(\Delta) \to 0} \sum_\nu \sum_\mu X_{\nu,\mu} \lambda_{\nu,\mu-1} = 0,$$

irrespective of the sub-divisions Δ_ν.

Next we shall prove convergence of the third term of (5.30). For this purpose we consider two sub-divisions Δ_ν, Δ'_ν of $(t_{\nu-1}, t_\nu)$. Then with suitably chosen values v_μ, $w_\mu \leqq \tau_\mu$ belonging to $(t_{\nu-1}, t_\nu)$ we can write

$$(5.34) \qquad \begin{cases} S_{\Delta_\nu} - S_{\Delta'_\nu} = \sum_\mu \dfrac{b'_{\mu-1} - b_{\mu-1}}{b_{\mu-1} b'_{\mu-1}} (\xi_\mu - \xi_{\mu-1}) \lambda_{\mu-1} \lambda'_{\mu-1} \\ + \sum_\mu \dfrac{b'_{\mu-1} - b_{\mu-1}}{b_{\mu-1} b'_{\mu-1}} (\xi_\mu - \xi_{\mu-1})(1 - \lambda_{\mu-1} \lambda'_{\mu-1}), \end{cases}$$

where

$$b_\mu = b[v_\mu, y(v_\mu)], \quad b'_\mu = b[w_\mu, y(w_\mu)], \quad \xi_\mu = \xi(\tau_\mu),$$
$$\lambda_\mu = \lambda[y(v_\mu)], \quad \lambda'_\mu = \lambda[y(w_\mu)].$$

Then, using the same arguments as in (5.31) and (5.32) for two summations in (5.34), we have

$$\Pr.\{|S_{\Delta_\nu} - S_{\Delta'_\nu}| > 2\delta\} \leqslant \delta^{-2} c(M)[\rho(\Delta_\nu) + \rho(\Delta'_\nu)]\Delta t_{\nu-1} + \varepsilon(M),$$

and so there exists

$$\text{(5.35)} \qquad \operatorname*{st.\,lim}_{\rho(\Delta_\nu)\to 0} S_{\Delta_\nu}.$$

But, in the same way we can show that

$$\Pr.\left\{\left|\sum_{\mu=1}^{n_\nu} \frac{\int_{u_{\nu,\mu-1}}^{u_{\nu,\mu}} a(\tau, y)\,d\tau - a[u_{\nu,\mu-1}, y(u_{\nu,\mu-1})](u_{\nu,\mu} - u_{\nu,\mu-1})}{b(u_{\nu,\mu-1}, y(u_{\nu,\mu-1}))}\right| > \delta\right\} \to 0,$$

$$\rho(\Delta_\nu) \to 0.$$

Therefore, in view of the results in the first step, the limit (5.35) is nothing but an increment $x(t_\nu) - x(t_{\nu-1})$ of the Brownian motion introduced there. Now making $\rho(\Delta_\nu) \to 0$, and then $\rho(\Delta) \to 0$ we get from (5.30), for every fixed t as a limit in probability

$$y(t) = y_0 + \int_0^t a(\tau, y)\,d\tau + \int_0^t b(\tau, y)\,dx(\tau), \qquad \text{q.e.d.}$$

It should be noted that, after an easy argument, as in the final step of the proof of Theorem 1, we might give the integral signs the precise meaning defined by Itō. In particular $y(t)$ is then continuous with probability 1. Indeed we have

THEOREM 8. *Suppose that* $a(t, x)$, $b(t, x)$ *are continuous and satisfy the Lipschitz condition* (1.1). *Then the solution of* (1.2.) *satisfies*

$$\overline{\lim_{h\to+0}} \frac{|y(t+h) - y(t)|}{\sqrt{2h\log\log h^{-1}}} = b(t, y(t))$$

with probability 1.

Proof. In the expression

$$y(t+h) = y(t) + \int_t^{t+h} a(\tau, y)\,d\tau + \int_t^{t+h} b(\tau, y)\,dx(\tau)$$

the first integral is an ordinary Riemann integral, with probability 1, of a continuous function. Hence with probability 1

$$\text{(5.36)} \qquad \frac{\int_t^{t+h} a(\tau, y)\,d\tau}{\sqrt{2h\log\log h^{-1}}} \to 0, \quad h \to +0.$$

If we write $\mathrm{T}(h) = \sup\limits_{0 \leq \sigma \leq h} \left| \int_t^{t+\sigma} [b(\tau, y(\tau)) - b(\tau, y(t))] dx(\tau) \right|$, then by a generalization of Kolmogorov's inequality (see Itō [9])

$$\mathrm{Pr.}\ \{\mathrm{T}(h) > \alpha\} \leqslant \alpha^{-2} \int_t^{t+h} E[b(\tau, y(\tau)) - b(\tau, y(t))]^2 d\tau$$

$$\leqslant c\alpha^{-2} \int_t^{t+h} (\tau - t) d\tau = \frac{ch^2}{2\alpha^2},$$

whence

$$\sum_{n=1}^{\infty} \mathrm{Pr.}\ \{\mathrm{T}(n^{-2}) > \eta \sqrt{n^{-2} \log\log n^2}\} \leqslant \sum_1^{\infty} \frac{cn^2 \cdot n^{-4}}{2\eta^2 \log\log n^2} < \infty$$

for every $\eta > 0$. We have thus proved that with probability 1

$$\lim_{n \to \infty} \mathrm{T}(n^{-2}) / \sqrt{n^{-2} \log\log n^2} = 0.$$

Suppose now we have chosen the integer n such that $(n+1)^{-2} < h \leqslant n^{-2}$, then with probability 1

$$(5.37) \qquad \left\{ \begin{aligned} & \left| \int_t^{t+h} [b(\tau, y(\tau)) - b(\tau, y(t))] dx(\tau) \right| \Big/ \sqrt{2h \log\log h^{-1}} \\ & \qquad \leqslant \mathrm{T}(n^{-2}) / \sqrt{2(n+1)^{-2} \log\log (n+1)^2} \to 0. \end{aligned} \right.$$

But, by the ordinary law of the iterated logarithm on the Brownian motion, with probability 1

$$(5.38) \qquad \overline{\lim_{h \to +0}} \frac{\left| \int_t^{t+h} b(\tau, y(t)) dx(\tau) \right|}{\sqrt{2h \log\log h^{-1}}} = b(t, y(t)).$$

Combination of (5.36), (5.37), (5.38) proves the theorem.

In the above theorem the passage to the limit of h is restricted to the one side, and a, b are supposed to satisfy the Lipschitz condition. It is left open to remove these restrictions.

In conclusion we will state an

EXAMPLE. Consider the well-known Gaussian stationary Markov process $y(t)$ with the auto-correlation $\rho(\tau) = e^{-\varrho|\tau|}$, $\rho > 0$. $y(t)$ is known to be represented as

$$y(t) = \sqrt{2\rho} \int_{-\infty}^t e^{-\varrho(t-\tau)} dx(\tau).$$

Then

$$dy(t) = \sqrt{2\rho}\,(-\rho e^{-\varrho t})\,dt \int_{-\infty}^{t} e^{\varrho\tau}\,dx(\tau) + \sqrt{2\rho}\,e^{-t\varrho} e^{\varrho t}\,dx(t)$$

$$= -\rho y(t)\,dt + \sqrt{2\rho}\,dx(t).$$

If we transform $y(t)$ into $z(t)$ by means of

$$z = \arctan y,$$

$z\,(t)$ satisfies

$$dz(t) = -\rho\left(\frac{y}{1+y^2} + \frac{2y}{(1+y^2)^2}\right)dt + \frac{\sqrt{2\rho}}{1+y^2}\,dx(t),$$

or

$$dz(t) = -\rho \cos z \sin z\,(1+\cos^2 z)\,dt + \sqrt{2\rho}\cos^2 z\,dx(t).$$

The formal arguments used here will be easily justified. Thus obtained Markov process $z(t)$ is also stationary, but non-gaussian, with the stochastic equation having periodic coefficients. This fact corresponds to that the phase space $(-\infty, \infty)$ consists of the ergodic parts $(-\pi/2 + n\pi,\ \pi/2 + n\pi)$, $n = 0, \pm 1, \pm 2, \dots$.

Tokio, October 1953

REFERENCES

1. S. Bernstein, *Équations différentielles stochastiques*, Actualités Scientifiques et Industrielles, n. 738, Herman, Paris, 1938, pp. 5-32.
2. M. D. Donsker, *An invariance principle for certain probability limit theorems*, Memoirs of the American Mathematical Society, Vol. 5 (1951), pp. 43-55.
3. J. L. Doob, *Heuristic approach to the Kolmogorov-Smirnov theorems*, Annals of Mathematical Statistics, Vol. 20 (1949), pp. 394-403.
4. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.
5. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bulletin of the American Mathematical Society, Vol. 52 (1946), pp. 292-302.
6. W. Feller, *Zur Theorie der stochastischen Prozesse*, Mathematische Annalen, Vol. 113. (1936), pp. 113-160.
7. W. Feller, *On the integrodifferential equations of purely discontinuous Markoff processes*, Transactions of the American Mathematical Society, Vol. 48 (1940), pp. 488-515.
8. R. Fortet, *Les fonctions aléatoires du type de Markoff associées à certaines équations linéaires aux derivées partielles du type parabolique*, Journal de Mathematiques pures et appliquées, Vol. 22 (1943), pp. 177-243.
9. K. Itō, *Stochastic integral*, Proceedings of the Imperial Academy, Tokyo, Vol. 20 (1944), pp. 519-524.
10. K. Itō, *On a stochastic integral equation*, Proceedings of Japan Academy, Vol. 22 (1946) pp. 32-35.
11. K. Itō, *On stochastic differential equations*, Memoirs of the American Mathematical Society, Vol. 4 (1951), 51 pp.
12. M. Kac, *On some connections between probability theory and differential and integral equations*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 1951, pp. 189-215.
13. A. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933.
14. A. Kolmogoroff, *Ueber die analytischen Methoden der Wahrscheinlichkeitsrechnung*, Mathematische Annalen, Vol. 104 (1931), pp. 415-458.
15. G. Maruyama, *Markov proceesses and stochastic equations*, Natural Science Report, Ochanomizu University, Vol. 4 (1953), pp. 40-43.
16. M. Udagawa, *Asymptotic properties of distributions of some functionals of random variables*, Reports of Statistical Application Research, Union of Japanese Scientists and Engineers, Vol. 2 (1952), pp. 1-66.