MEMOIRS
OF THE
AMERICAN MATHEMATICAL SOCIETY

NUMBER 4

ON STOCHASTIC DIFFERENTIAL EQUATIONS

BY

KIYOSI ITO

PUBLISHED BY THE
AMERICAN MATHEMATICAL SOCIETY
531 West 116th St., New York City
1951

ON STOCHASTIC DIFFERENTIAL EQUATIONS

By

KIYOSI ITO

Let x_t be a simple Markoff process with a continuous parameter t, and $F(t,\xi;s,E)$ be the transition probability law of the process:

(1) $F(t,\xi;s,E) = Pr\{x_s \in E/x_t = \xi\}$,

where the right side means the probability of $x_s \in E$ under the condition: $x_t = \xi$. The differential of x_t at $t = s$ is given by the transition probability law of x_t in an infinitesimal neighborhood of $t = s$:

(2) $F(s-\Delta_2, \xi; s+\Delta_1, E)$.

W. Feller[1] has discussed the case in which it has the following form:

(3) $F(s-\Delta_2, \xi; s+\Delta_1, E) = (1-p(s,\xi)\ (\Delta_1+\Delta_2)G(s-\Delta_2, \xi; s+\Delta_1, E)$

$+(\Delta_1+\Delta_2)p(s,\xi)P(s,\xi,E) + o(\Delta_1+\Delta_2)$,

where $G(s-\Delta_2, \xi; s+\Delta_1, E)$ is a probability distribution as a function of E and satisfies

(4) $\frac{1}{\Delta_1+\Delta_2} \int_{|\eta-\xi|>\delta} G(s-\Delta_2, \xi; s+\Delta_1, d\eta) \longrightarrow 0$,

(5) $\frac{1}{\Delta_1+\Delta_2} \int_{|\eta-\xi|\le\delta} (\eta-\xi)^2\, G(s-\Delta_2, \xi; s+\Delta_1, d\eta) \longrightarrow 2a(t,\xi)$,

(6) $\frac{1}{\Delta_1+\Delta_2} \int_{|\eta-\xi|\le\delta} (\eta-\xi) G(s-\Delta_2, \xi; s+\Delta_1, d\eta) \longrightarrow b(t,\xi)$,

for $\Delta_1+\Delta_2 \longrightarrow 0$ and $p(s,\xi) \geq 0$ and $P(s,\xi,E)$ is a probability distribution in E. The special case of "$p(s,\xi) = 0$" has already been treated by A. Kolmogoroff[2] and S. Bernstein.[3]

We shall introduce a somewhat general definition of the differential of the process x_t (Cf. §5). Let $P_{s,\xi,\Delta_1,\Delta_2}$ denote the conditional probability law:

$Pr\{x_{s+\Delta_1} - x_{s-\Delta_2} \in E/x_{s-\Delta_2} = \xi\}$, $\Delta_1, \Delta_2 \geq 0$.

If the $[1/\Delta_1+\Delta_2]$- times[4] convolution of $P_{s,\xi,\Delta_1,\Delta_2}$ tends to a probability law $L_{s,\xi}$ with regard to Lévy's law-distance as $\Delta_1+\Delta_2 \longrightarrow 0$, then $L_{s,\xi}$ is called the stochastic differential coefficient at s. $L_{s,\xi}$ is clearly an infinitely divisible law. In the above Feller's case the logarithmic characteristic function [5]

Received by the editors March 29, 1949.

$\psi(z, L_{s,\xi})$ of $L_{s,\xi}$ is given by

(7) $$(z, L_{s,\xi}) = ib(s,\xi)z - a(s,\xi)z^2 + p(s,\xi)\int_{-\infty}^{\infty}(e^{iuz}-1)P(s,\xi,du(+)\xi).^{6)}$$

A problem of stochastic differential equations is to construct a Markoff process whose stochastic differential coefficient $L_{t,\xi}$ is given as a function of (t,ξ). W. Feller has deduced the following integro-differential equation from (3), (4), (5) and (6):

(8) $$\frac{\partial}{\partial t}F(t,\xi;s,E) + a(t,\xi)\frac{\partial^2}{\partial \xi^2}F(t,\xi;s,E) + b(t,\xi)\frac{\partial}{\partial \xi}F(t,\xi;s,E)$$
$$-p(t,\xi)F(t,\xi;s,E) + p(t,\xi)\int_{-\infty}^{\infty}F(t,\eta;s,E)P(t,\xi,d\eta) = 0.$$
He has proved the existence and uniqueness of the solution of this equation under some conditions and has shown that the solution becomes a transition probability law, and satisfies (3),(4),(5) (6). He has termed the case: $p(t,\xi) \equiv 0$ as continuous case and the case: $a(t,\xi) \equiv 0$ and $b(t,\xi) \equiv 0$ as purely discontinuous case.

It is true that we can construct a simple Markoff process from the transition probability law by introducing a probability distribution into the functional space R^R by Kolmogoroff's theorem,[7] but it is impossible to discuss the regularity of the obtained process, for example measurability, continuity, discontinuity of the first kind etc., as was pointed out by J. L. Doob.[8] To discuss the measurability of the process for example, J. L. Doob has introduced a probability distribution on a subspace of R^R and E. Slutsky has introduced a new concept "measurable kernel".[9] We shall investigate the sense of the term "continuous case" and "purely discontinuous case" used by W. Feller from the rigorous view-point of J. L. Doob and E. Slutsky. A recent research of J. L. Doob[10] concerning a simple Markoff process taking values in an enumerable set has been achieved from this view-point. A research of R. Fortet[11] concerning the above continuous case seems also to stand on the same idea but the author is not yet informed of the details.

In his paper "ON STOCHASTIC PROCESSES (I)"[12] the author has deduced Lévy's canonical form of differential processes with no fixed discontinuities by making use of the rigorous scheme of J. L. Doob. Using the results of the above paper, we shall here construct the solution of the above stochastic differential equation in such a way that we may be able to discuss the regularity of the solution. For this purpose we transform the stochastic differential equation into a stochastic integral equation.

The first and most simple form of stochastic integral is Wiener's integral[13] which is an integral of a function $\sigma(t) \in L_2$ based on a brownian motion $g(t)$: $\int \sigma(t)dg(t)$. In this integral $\sigma(t)$ is not a random function. The author has ex-

extended this notion and defined an integral in case $\sigma(t)$ is a random function satissying some conditions.[14] A brownian motion is a temporally homogeneous and differential (i.e. spatially homogeneous) process with no moving discontinuity. The process $x(t) = \int_a^t \sigma(t)dg(t)$ obtained by Wiener's integral is not temporally homogeneous but spatially homogeneous. In order to obtain a simple Markoff process--which is in general neither temporally nor spatially homogeneous--we shall have to solve a stochastic integral equation:

$$x(t) = \int_a^t \sigma(\tau, x(\tau))dg(\tau)$$

or more generally

$$x(t) = c + \int_a^t m(\tau, x(\tau))d\tau + \int_a^t \sigma(\tau, x(\tau))dg(\tau).$$

The author has published a note[15] on this stochastic integral equation, which concerns the continuous case above mentioned.

In order to discuss the general case we shall have to consider a stochastic integral equation where the integral is based not on a brownian motion but on a more general temporally homogeneous differential process, which will be called a fundamental differential process (Cf. § 6) in this paper.

Chapter I is devoted to the explanation of the fundamental concepts. Some of them are well-known but we shall explain them in a rigorous form for the later use. In Chapter II we shall introduce a stochastic integral of a general type. The results of the author's previous paper[16] will be contained here in an improved form. The aim of this paper will be attained in Chapter III, where we shall investigate a stochastic differential equation and a stochastic integral equation.

The author expresses his hearty thanks to Professor S. Iyanaga, Professor K. Yosida, Professor S. Kakutani and Mr. H. Anzai who have encouraged him with their kind discussions and to Professor J. L. Doob who has given him valuable suggestions to improve the manuscript and friendly aid to publish it.

I. Fundamental concepts.

§1. Function of random variables. Let X be any set and B_X be a completely additive class of subsets of X. When we consider X together with B_X we call it a Borel field (X, B_X). It is evident that B_X may be arbitrarily taken, but in case X is the real number space R^1, then we usually take the system B^1 of all Borel subsets of R^1 as B_X, and in case X is R^A, B_X is usually the least completely additive class that contains all Borel cyclinder subsets of R^A, which we denote by B^A. If B_X and B_Y are associated respectively with X and with Y, then we usually associate with the product space the least completely additive class that contains all the sets of the form: $E_X \otimes Y$, $X \otimes E_Y$, $E_X \in B_X$, $E_Y \in B_Y$; this class will be denoted by $B_X \otimes B_Y$. The product of many Borel fields can be similarly defined.

Let (X,B_X) and (Y,B_Y) be Borel fields. A mapping $f(x)$ from X into Y is called to be B-measurable if $f^{-1}(E_Y) \in B_X$ for any $E_Y \in B_Y$. If $f(x)$ is a B-measurable mapping from (X,B_X) into (Y,B_Y) and if $g(x)$ is a B-measurable mapping from (Y,B_Y) into (Z,B_Z), then $g(f(x))$ will be a B-measurable mapping from (X,B_X) into (Z,B_Z).

Let (Ω,B_Ω,P) be a probability field, where Ω is a set, B_Ω is a completely additive class of subsets of Ω, and P is a probability distribution (p.d.) on (Ω,B_Ω). An (X,B_X)-valued function $x(\omega)$ defined on Ω is called an (X,B_X)-valued random variable, if it is B-measurable i.e. $x^{-1}(E_X) \in B_\Omega$ for any $E_X \in B_X$. If we put $P_x(E_X) = P(x^{-1}(E_X))$ for $E_X \in B_X$. P_x is a p.d. on a Borel field (X,B_X) which is called the probability law (P.ℓ.) of x; we also say that x is governed by P_x.

Let $x(\omega)$ be an (X,B_X)-valued random variable and $f(.)$ be a B-measurable mapping from (X,B_X) into (Y,B_Y). Put $y(\omega) = f(x(\omega))$. Then $y(\omega)$ will be a (Y,B_Y)-valued random variable. $y(\omega)$ is called a B-measurable function of $x(\omega)$.

Theorem 1. Let $y_n(\omega)$, $n=1,2,\ldots$, be real-valued B-measurable functions of an (X,B_X)-valued random variable. If $y_n(\omega)$ be convergent in probability, then the limit variable $y(\omega)$ is also coincident with a B-measurable function of $x(\omega)$ up to P-measure 0.

Proof. By taking a subsequence if necessary, we may assume that $y_n(\omega)$ be convergent with P-measure 1. Put $y_n(\omega) = f_n(x(\omega))$. Then

$$P_x\Big(\bigcap_p \bigcup_q \bigcap_{m,n>q} \{\xi;|f_m(\xi) - f_n(\xi)| < 1/p\}\Big)=P\Big(\bigcap_p \bigcup_q \bigcap_{m,n>q} \{\omega;|f_m(x(\omega)) - f_n(x(\omega))| < 1/p\}\Big) = 1.$$

Put $f(\xi) = \lim f_n(\xi)$ in the above ξ-set and $= 0$ elsewhere. Then $f(\xi)$ is a B-measurable function of $\xi \in (X,B\ )$, since the above ξ-set belongs to B_X. We have clearly, with probability 1,

$$f(x(\omega)) = \lim_n f_n(x(\omega)) = \lim_n y_n(\omega) = y(\omega),$$

which completes the proof.

§2. Conditional probability law. Let $x(\omega)$ and $y(\omega)$ be random variables taking values in (X,B_X) and (Y,B_Y) respectively. A function $P_y(E_Y/\xi)$ of $E_Y \in B_Y$ and $\xi \in X$ will be called the (conditional) probability law of $y(\omega)$ under the condition that $x(\omega) = \xi$ and will be denoted by $P_y(E_Y/x(\omega) = \xi)$ or $Pr\{y \in E_Y/x = \xi\}$, if and only if

(2.1) $P_y(E_Y/\xi)$ is a p.d. on (Y,B_Y) for any ξ,

(2.2) $P_y(E_Y/\xi)$ is a B-measurable function of $\xi \in (X,B)$ for any $E_Y \in B_Y$, and

(2.3) $\int_{E_X} P_y(E_Y/\xi) P_x(d\xi) = Pr\{x \in E_X \,\&\, y \in E_Y\} = P(x^{-1}(E_X) \cap y^{-1}(E_Y))$.

The existence and uniqueness (up to P-measure 0) of $P_y(E_Y/\xi)$ was proved by J. L. Doob[17] in the case that (Y,B_Y) is the n-dimensional space (R^n,B^n).

$P_y(E_Y/x(\omega))$ i.e. the function of ω obtained by replacing ξ with $x(\omega)$ in $P_y(E/\xi)$ will be called the conditional p.ℓ. of $y(\omega)$ under the condition that $x(\omega)$ is determined and it will be also denoted by $Pr\{y \in E_Y/x(\omega)\}$: this is clearly a real-valued random variable for any assigned E_Y. By (2.3) we have

(2.4) $\mathcal{E} P_y(E_Y/x(\omega)) = Pr\{y \in E_Y\}$ ($\mathcal{E}$ = expectation.).

If the p.ℓ. of the combined random variable $(x(\omega),y(\omega))$, which clearly takes values in $(X \otimes Y, P_X \otimes B_Y)$, is coincident with the direct product measure of P_x and P_y: $P_x \otimes P_y$ on $(X \otimes Y, B_X \otimes B_Y)$ then $x(\omega)$ and $y(\omega)$ are called to be independent. The independence of many random variables can be similarly defined. Clearly we have

Theorem 2.1. $x(\omega)$ and $y(\omega)$ be independent. Then

$P_y(E_Y/x(\omega) = \xi) = P_y(E_Y)$ for almost all (P_x) ξ,

i.e.

$P_y(E_Y/x(\omega)) = P_y(E_Y)$ for almost all (P) ω.

Theorem 2.2. $x(\omega)$ and $y(\omega)$ be independent. $G(\xi,\eta)$ be a B-measurable mapping from $(X \otimes Y, B_X \otimes B_Y)$ into (R^1,B^1). Put $z(\omega) = G(x(\omega),y(\omega))$. Then we have

$P_z(E/x(\omega) = \xi) = Pr\{G(\xi,y(\omega)) \in E\}$

for almost all (P_x) ξ.

Proof. Since $x(\omega)$ and $y(\omega)$ are independent, we can make use of Fubini's theorem.

$$\Pr\{z\in E, x\in E_X\} = (P_x \otimes P_y)(\{(\xi,\eta); f(\xi,\eta)\in E, \xi\in E_X\})$$

$$= \int_{E_X} P_y(\{\eta; f(\xi,\eta)\in E\})\, P_x(d\xi) = \int_{E_X} \Pr\{f(\xi, y(\omega))\in E\} P_x(d\xi),$$

which completes the proof.

Theorem 2.3. $x(\omega)$ and $y(\omega)$ be independent. $G(\xi,\eta)$ be any real-valued B-measurable function in (ξ,η). If $G(x(\omega),y(\omega)) = 0$ with P-measure 1, then $G(\xi,y(\omega))=0$ with P-measure 1 for almost all (P_x) ξ.

Proof. By Theorem 2.2 we have $\int_X \Pr\{G(\xi,y(\omega)) = 0\} P_x(d\xi)\ \Pr\{G(x(\omega),y(\omega))=0\}=1$ and so $\Pr\{G(\xi,y(\omega)) = 0\} = 1$ for almost all (P_x) ξ.

§3. Transition probability law. $x(\tau,\omega)$ be a real random variable for any τ, $a \le \tau \le b$. The system $x(\tau,\omega)$, $a \le \tau \le b$, is called a stochastic process, which is also considered as an (R^I,B^I)-valued random variable, I being the interval $[a,b]$[18]. The p.l. of $x(s,\omega)$ under the condition that $(x(\tau,\omega),\ a \le \tau \le t)$[19] is determined:

(3.1) $\Pr\{x(s,\omega)\in E/x(\tau,\omega),\ a \le \tau \le t\}$ $(t < s)$

is called the transition probability law of this process. If this is equal to

(3.2) $\Pr\{x(s,\omega) \in E/x(t,\omega)\}$

for almost all (P) ω, the process is called a simple Markoff process. In such a process we put

(3.3) $F(t,\xi;s,E) = \Pr\{x(s,\omega) \in E/x(t,\omega) = \xi\}$.

Then we can easily prove, for almost all $(P_{x(t,\omega)})$ ξ,

(3.4) $F(t,\xi;s,E) = \int_{-\infty}^{\infty} F(t,\xi;u,d\eta)F(u,\eta;s,E),$ $(t < u < s)$,

which is well-known as Chapman's equation.

If $x(s_\nu,\omega) - x(t_\nu,\omega)$, $\nu=1,2,\ldots,n$, are independent random variables for any system of non-overlapping intervals (t_ν,s_ν), $\nu=1,2,\ldots,n$, then we call $x(\tau,\omega)$, $a \le \tau \le b$, a differential process. This is evidently a simple Markoff process whose transition p.l. is given by

(3.5) $F(t,\xi;s,E) = F_{t,s}(E(-)\xi)$,

where $F_{t,s}$ is the p.l. of $x(s,\omega)-x(t,\omega)$ and $E(-)\xi$ is the set $\{\eta - \xi; \eta\in E\}$; (3.5) will be obtained at once if we substitute $(x(\tau,\omega),\ a \le \tau \le t)$, $x(s,\omega)-x(t,\omega)$ and $x(s,\omega)$ respectively for $x(\omega)$, $y(\omega)$ and $z(\omega)$ in Theorem 2.2.

§4. THREE ELEMENTS OF AN INFINITELY DIVISIBLE LAW OF PROBABILITY.

The logarithmic characteristic function (l.c.f.) of an infinitely divisible law of probability (i.d.l.) can be expressed in the form:

$$(4.1)\quad imz - \frac{\sigma^2}{2} z^2 + \int_{|u|>1} (e^{if(u)z} - 1) \frac{du}{u^2} + \int_{|u|\leq 1} (e^{if(u)z} - 1 - if(u)z) \frac{du}{u^2}$$

in one and only one way, where m is real, $\sigma \geq 0$, and $f(u)$ is monotone non-decreasing and right-continuous and

$$\int_{|u|\leq 1} f(u)^2 \frac{du}{u^2} < \infty ;$$

this formula is deduced at once from Levy's formula.[20] These $m, \sigma, f(u)$ will be called the <u>three elements</u> of this i.d.l. . The i.d.l. whose l.c.f. is

$$(4.2)\quad \Psi_0(z) = iz - \frac{z^2}{2} + \int_{|u|>1} (e^{iuz} - 1) \frac{du}{u^2} + \int_{|u|\leq 1} (e^{iuz} - 1 - izu) \frac{du}{u^2} ,$$

i.e. $m = \sigma = 1$, $f(u) \equiv u$,

will be called the fundamental i.d.l. in this note.

Theorem 4.1. Let $m(L), \sigma(L)$ and $f(u,L)$ be the three elements of an i.d.l. L. Then $m(L), \sigma(L)$ and $f(u,L)$ (for any fixed u) are all B-measurable in $L = (L(E); E \in B^1) \in (R^{B^1}, B^{B^1})$

Remark. By the expression " $m(L)$ is B-measurable in $L \equiv (L(E), E \in B^1) \in (R^{B^1}, B^{B^1})$ we mean that there exists at least one B-measurable function $M(L)$ defined on the whole space (R^{B^1}, B^{B^1}) such that we have $M(L) = m(L)$ for any $L \equiv (L(E), E \in B^1)$ that is an i.d.l. as a function of E.

Proof. Let $\phi(z,L)$ be the characteristic function of any i.d.l. L. For any z, $\phi(z,L)$ is B-measurable in $L \in (R^{B^1}, B^{B^1})$, because, if we define $\Phi_z(L)$ by

$$\Phi_z(L) = \lim_{n\to\infty} \sum_{k=-n^2}^{n^2} \exp(ikz/n) L((k-1/n, k/n)) \quad \text{(if this limit exists)}$$
$$= 0 \quad \text{(if otherwise)},$$

then $\Phi_z(L)$ (for each z) is B-measurable function defined on the whole space (R^{B^1}, B^{B^1}) and $\Phi_z(L) = \phi(z,L)$ for any i.d.l. L.

Let $\psi(z,L)$ be the logarithmic characteristic function of any i.d.l. Since $\psi(z,L)$ is the branch of log $\phi(z,L)$ which is obtained from log $\phi(0,L)=1$ by the analytic prolongation along the curve:

$$\phi(\lambda,L), \quad o \leq \lambda \leq z \text{ (or } o \geq \lambda \geq z)$$

and so it is expressible as

$$\psi(z,L) = \lim_{n\to\infty} \sum_{k=1}^{n} \int_0^1 \frac{\phi(\frac{k}{n}z,L) - \phi(\frac{k-1}{n}z,L)}{(\phi(\frac{k}{n}z,L) - \phi(\frac{k-1}{n}z,L))t+1}\, dt$$

we see that $\psi(z,L)$ is also B-measurable in L for any z. By virtue of the Levy's formula $\psi(z,L)$ is written in the form

$$\psi(z,L) = i\bar{m}(L)z - \frac{\sigma^2(L)}{2}z^2 + \int_{-\infty}^{\infty}(e^{izu} - 1 - \frac{izu}{1+u^2})n(du,L),$$

where the measure n is determined by the following procedure (Cf. A. Khintchine: Déduction nouvelle d'une formule de P. Lévy, Bull. d. l'univ. d'etat a Moscou, Serie International, Sect. A, Vol. 1, Fasc. 1, 1937),

$$\Delta(t,L) = \int_{t-1}^{t+1} \psi(z,L)dz - 2\psi(t,L),$$

$$K(u,L) = \frac{1}{2\pi} \lim_{c\to\infty} \int_{-c}^{c} \frac{1-e^{itu}}{it} \Delta(t,L)dt,$$

$$G(v,L) = -\int_{-\infty}^{v} \frac{1}{2(1-\frac{\sin u}{u})} dK(u,L),$$

$$n((a,\infty), L) = \int_{a-o}^{\infty} \frac{1+v^2}{v^2} dG(v,L) \qquad (a > 0),$$

$$n((-\infty,a),L) = \int_{-\infty}^{a+o} \frac{1+v^2}{v^2} dG(v,L) \quad (a < 0).$$

Therefore we can prove recursively the B-measurability of the above functions of L. Thus we obtain, for each $a > 0$, a B-measurable functions $N_a(L)$ defined on the whole space (R^{B^1}, B^{B^1}) such that $N_a(L) = n((a,\infty),L)$ for any i.d.l. L. We may assume that $N_a(L)$ is monotone-decreasing and left-continuous in a for each L, by taking the supremum of

$N_r(L)$, r running over all rational numbers $r < a$, instead of $N_a(L)$, if necessary.

Now we shall prove, for each $u > 0$, that $f(u,L)$ is B-measurable in L. $f(u,L)$ is written in the following form by the definition.

$$f(u,L) = \inf \{a; n((a,\infty),L) < \frac{1}{u}\} \quad (u > 0).$$

Therefore, if we put

$$F_u(L) = \inf\{a; N_a(L) < \frac{1}{u}\},$$

$F_u(L)$ (for each $u > 0$) is a function defined on the whole space RB^1 and $F_u(L)=f(u,L)$ for any i.d.l. L. The B-measurability of $F_u(L)$ is clear on account of the fact that "$F_u(L) < a$" is equivalent to "$N_a(L) < \frac{1}{u}$", which follows from the definition of $F_u(L)$ and the monotone-property (in a) of $N_a(L)$. Thus we see that $f(u,L)$ (for each $u > 0$) is B-measurable in L. Similarly we can show that $f(u,L)$ $(u < 0)$ is B-measurable in L. It is clear that $f(0,L)$ $(\equiv 0)$ is B-measurable in L.

Now we put

$$\Phi(z,L) \equiv im(L)z - \frac{\sigma^2(L)}{2}z^2 \equiv \psi(z,L) - \int_{|u|>1} (\exp(izf(u,L))-1) \frac{du}{u^2}$$

$$- \int_{|u|\leq 1} (\exp(izf(u))-1-izf(u,L)) \frac{du}{u^2} .$$

Then $\Phi(z,L)$ (for each z) is B-measurable in L, since $\psi(z,L)$ and $f(u,L)$ are B-measurable. But we have

$$m(L) = \frac{1}{2i}(4\Phi(1,L) - \Phi(2,L))$$

and

$$\sigma^2(L) = 2\Phi(1,L) - 4\Phi(2,L),$$

from which follows the B-measurability of $m(L)$ and $\sigma(L)$.

Theorem 4.2. Let L_α, $\alpha \in A$, be any system of i.d.l. depending on $\alpha \in A$ and m_α, σ_α and $f_\alpha(u)$ be the three elements of L_α. In order that L_α, $\alpha \in A$, be totally bounded in the sense of Levy's law-distance,[21] it is necessary and sufficient that each of $|m_\alpha|$, σ_α and $||f_\alpha||_n$, n=1,2,..., is bounded, where

$$||f_\alpha||_n^2 = \int_{|u|\leq n} f_\alpha(u)^2 \frac{du}{u^2} .$$

Proof. L_α is decomposed as

$$L_\alpha = L^{(1)}_\alpha * L^{(2)}_\alpha * L^{(3)}_{\alpha n} * L^{(4)}_{\alpha n} * L^{(5)}_{\alpha n},$$

where the l.c.f. of the factors are respectively

$$im_\alpha z,\ -\frac{\sigma_\alpha^2}{2} z^2,\ iz\int_{1<|u|<n} f_\alpha(u)\frac{du}{u^2},\ \int_{o<|u|<n}(e^{if_\alpha(u)z}-1-if_\alpha(u)z)\frac{du}{u^2}$$

and $\displaystyle\int_{|u|\geq n}(e^{if_\alpha(u)z}-1)\frac{du}{u^2}$.

Sufficiency. If the condition is satisfied, $\{L^{(1)}_\alpha\}$, $\{L^{(2)}_\alpha\}$ are clearly totally bounded and $\{L^{(3)}_{\alpha n}\}$ is also totally bounded for any fixed n, since we have, by Schwarz' inequality,

$$\left|\int_{1<|u|<n} f_\alpha(u)\frac{du}{u^2}\right|^2 \leq 2\int_{1<|u|<n} f_\alpha(u)^2\frac{du}{u^2}.$$

$L^{(4)}_{\alpha n}$ has the expectation 0 and the standard deviation $||f_\alpha||$ and so $\{L^{(4)}_{\alpha n}, \alpha\in A\}$ is totally bounded. Therefore

$$\{L^*_{\alpha n} \equiv L^{(1)}_\alpha * L^{(2)}_\alpha * L^{(3)}_{\alpha n} * L^{(4)}_{\alpha n}, \alpha\in A\}$$

is totally bounded, and so we have

$$\lim_{c\to\infty} \inf_\alpha L^*_{\alpha n}((-c,c)) = 1.$$

But

$$L_\alpha((-c,c)) \geq L^*_{\alpha n}((-c,c))\, L^{(5)}_\alpha(\{0\}) = L^*_{\alpha n}((-c,c))\exp(-2/n).$$

Consequently we have

$$\lim_{c\to\infty} \inf_\alpha L_\alpha((-c,c)) \geq \exp(-2/n) \text{ and so } \lim_{c\to\infty} \inf_\alpha L_\alpha((-c,c))=1,$$

which completes the proof.

Necessity. Let $Q(L,c)$ be Levy's concentration function[22] of the p.d.L. Suppose that $\{L_\alpha\}$ is totally bounded. Then

$$\inf_\alpha Q(L^{(2)}_\alpha, c) \longrightarrow 1 \qquad \text{as } c\longrightarrow\infty.$$

But we have $Q(L^{(2)}_\alpha, c) \geq Q(L_\alpha, c)$ by Levy's theorem concerning the non-decreasing of concentration function. Therefore

$$\inf_\alpha Q(L_\alpha^{(2)},c) \longrightarrow 1 \quad \text{as} \quad c \longrightarrow \infty,$$

and so σ_α will be bounded since $L_\alpha^{(2)}$ is a Gaussian distribution with the mean 0 and the standard deviation σ_α.

$L^{(5)}_{\alpha\, n}$ is decomposed as $L^{(5)}_{\alpha\, n} = L^{(5)}_{\alpha\, n+} * L^{(5)}_{\alpha\, n-}$, where the factors has the l.c.f.

$$\int_n^\infty (\exp(if_\alpha(u)z) - 1)du/u^2 \qquad \text{and} \qquad \int_{-\infty}^{-n} (\exp(if_\alpha(u)z) - 1)du/u^2$$

respectively. By the above-cited Lévy's theorem we see

(4.3) $\inf_\alpha Q(L^{(5)}_{\alpha\, n+},c) \geq \inf_\alpha Q(L_\alpha,c) \longrightarrow 1$ as $c \longrightarrow \infty$.

But $c < f_\alpha(n)$ implies $Q(L^{(5)}_{\alpha\, n+},c) = \exp(-1/n)$, i.e.

(4.4) $Q(L^{(5)}_{\alpha\, n},c) > \exp(-1/n)$ implies $c \geq f_\alpha(u)$.

By (4.3) there exists c such that $Q(L^{(5)}_{\alpha\, n+},c) > \exp(-1/n)$ and so that $c \geq f_\alpha(n)$ for $\alpha \in A$. Thus we see that $f_\alpha(n)$ is bounded for any assigned n. This is also the case for $f_\alpha(-n)$. Consequently we see that $f_\alpha(u)$ is bounded whenever $\alpha \in A$ and $|u| \leq n$, for any fixed n.

If $L_{\alpha(p)}$, p=1,2,..., be chosen from $\{L_\alpha\}$ such that $||f_{\alpha(p)}||_n$ increases indefinitely with p, $L^{(4)}_{\alpha(p)n}$ is approximately a Gaussian distribution with the mean 0 and the standard deviation $||f_{\alpha(p)}||$ as $p \longrightarrow \infty$ by the central limit theorem. Thus we have

$$Q(L^{(4)}_{\alpha(p)n}, ||f_{\alpha(p)}||_n) \longrightarrow \int_{-1}^{1} 1/\sqrt{2\pi}\, \exp(-t^2/2)dt < 1$$

as $p \longrightarrow \infty$, which contradicts with the fact that

$$\inf_\alpha Q(L^{(4)}_{\alpha\, n},c) \geq \inf_\alpha Q(L_\alpha,c) \longrightarrow 1 \ (\text{as } p \longrightarrow \infty).$$

Thus $||f_\alpha||_n$ proves to be bounded for any fixed n. Therefore

$\{L^{(3)}_{\alpha\, n} * L^{(4)}_{\alpha\, n} * L^{(5)}_{\alpha\, n}\}$ is totally bounded. Therefore $L_\alpha^{(1)}$ must be totally bounded and so $\{m_\alpha\}$ will be bounded.

§5. STOCHASTIC DIFFERENTIATION. $x(\tau,\omega)$, $a \leq \tau \leq b$, be a stochastic process on (Ω,B,P) and $F(E,\omega;\Delta_1,\Delta_2)$ be the conditional p.l. of $x(t+\Delta_1,\omega)-x(t-\Delta_2,\omega)(\Delta_1,\Delta_2 \geq 0)$ under the condition that $x(\tau,\omega)$, $a \leq \tau \leq t-\Delta_2$, be determined. If the $[1/\Delta_1+\Delta_2]$-times convolution of $F(E,\omega;\Delta_1,\Delta_2)$ ρ-converges to a distribution $L(E,\omega)$ in probability as $\Delta_1+\Delta_2 \longrightarrow 0$, i.e. for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ such that $0 < \Delta_1+\Delta_2 < \delta$ implies

$$\Pr\{\rho(H(E,\omega;\Delta_1,\Delta_2), L(E,\omega)) > \varepsilon\} < \varepsilon,$$

ρ being Lévy's law-distance, then we say that $x(\tau,\omega)$ is differentiable at t and we call $L(E,\omega)$ the differential coefficient of $x(\tau,\omega)$ at t, and we denote it with $D_t x(\tau,\omega)$ or briefly with $D(t,\omega)$. This is considered as an (RB^1,BB^1)-valued random variable. By taking a convenient sequence $\Delta_1'+\Delta_2' > \Delta_1''+\Delta_2'' > \ldots \longrightarrow 0$, we see that $L(E,\omega)$ is the ρ-limit of the $[1/\Delta_1^{(n)}+\Delta_2^{(n)}]$-times convolution of $F(E,\omega;\Delta_1^{(n)},\Delta_2^{(n)})$ with P-measure 1, and so we obtain

Theorem 5.1. $Dx(t,\omega)$ is an i.d.l. with P-measure 1.

From the definition we obtain, by making use of Theorem 1,

Theorem 5.2. $Dx(t,\omega)$ is a B-measurable function of $(x(\tau,\omega), a \leq \tau \leq t)$. If $x(\tau,\omega)$, $a \leq \tau \leq b$, is a simple Markoff process, then $Dx(t,\omega)$, if it exists, is a B-measurable function of $x(t,\omega)$; the form of the function clearly depends on t. If $x(\tau,\omega)$, $a \leq \tau \leq b$, is a differential process, then $Dx(t,\omega)$, if it exists, does not depend on ω but on t. If $x(\tau,\omega)$, $a \leq \tau \leq b$, is a temporally homogeneous differential process, then $Dx(t,\omega)$ exists and depends neither on ω nor on t; the l.c.f. of $Dx(t,\omega)$ is equal to the l.c.f. of the p.l. of $x(b,\omega) - x(a,\omega)$ divided by b-a.

We can easily see that, if F_n^{*n} ρ-converges to a probability law, then F_n ρ-converges to the unit distribution, and so we have

Theorem 5.3. If $x(\tau,\omega)$ is differentiable at t, then it is continuous at t in probability i.e. t is not a fixed discontinuity of this process.

II. Stochastic Integral.

The integral of the form:

$$\int \sigma(\tau)\,dg(\tau,\omega),$$

where $\sigma(\tau) \in L_2$ and $g(\tau, \omega)$ is a brownian motion, is well-known as Wiener's integral.[23] The author has extended this integral to the case in which σ depends not only on τ but also on ω and called it a stochastic integral.[24] In this Chapter we treat a more general stochastic integral for the later use.

§6. FUNDAMENTAL DIFFERENTIAL PROCESS. Let $l(t,\omega)$, $a \leq t \leq b$, be a temporally homogeneous differential process such that both $l(t+0,\omega)$ and $l(t-0,\omega)$ exist and $l(t+0,\omega) = l(t,\omega)$, i.e. $l(t,\omega)$ is continuous in t except possibly for discontinuities of the first kind (hereafter we term this property with " belong to d_j-class"). Further we require that the p.l. of $l(s,\omega) - l(t,\omega)$ has the l.c.f. $(s-t)\psi_o(z)$, where $\psi_o(z)$ is the l.c.f. of the fundamental i.d.l. . Then $l(t,\omega)$, $a \leq t \leq b$, is defined to be a fundamental differential process. Such a process can be realized on a conveniently defined probability field (Ω, B_Ω, P), where the p.l. of $l(a,\omega)$ can be arbitrarily assigned.

Any jump of $l(t,\omega)$ is expressed by a point $(t,u) \in [a,b] \otimes R^1$, t being its position and u being its height: $l(t,\omega) - l(t-0,\omega)$. The number $p(E,\omega)$ of the jumps in E, E being a Borel subset of $[a,b] \otimes R^1$, can be considered a real random variable, which proves to be governed by the Poisson distribution with the mean:

$$\pi(E) = \int_E d\tau du/u^2.$$

$p(E,\omega)$ is evidently a function of $l(t,\omega)$, $a \leq t \leq b$. The system $\{p(E,\omega)\}$ is called the discontinuous part of $l(t,\omega)$, $a \leq t \leq b$, $l(t,\omega)$ can be expressed as

$$l(t,\omega) = l(a,\omega) + t + g(t,\omega) + \int_a^t \int_{|u|>1} up(d\tau du,\omega) + \int_a^t \int_{|u|<1} uq(d\tau du,\omega)$$

for any t, $a \leq t \leq b$, for almost all (P) ω, where $q(E,\omega) = p(E,\omega) - \pi(E)$ and $g(t,\omega)$ is a brownian motion which is also a function of $l(\tau,\omega)$, $a \leq \tau \leq t$, and is called the continuous part of $l(\tau,\omega)$.

For any disjoint system $E_1, E_2, \ldots, E_n$, $p(E_1,\omega), p(E_2,\omega), \ldots, p(E_n,\omega)$ and $(g(\tau,\omega), a \leq \tau \leq b)$ are independent.

All these properties can be immediately deduced from the results in the above-cited paper.[25]

§7. STOCHASTIC INTEGRAL BASED ON g. We shall define here an integral of the form:

(7.1) $\int_E \sigma(\tau,\omega)dg(\tau,\omega)$, E being a Borel subset of $(a,b]$, in such a way that it may be a natural extension of Wiener's integral.

First we shall consider the case in which E is an interval: $I_1=(\alpha,\beta]$. By $S(I_1)$ we denote the class of all functions $\sigma(\tau,\omega)$, $\alpha\leq\tau\leq\beta$, $\omega\in\Omega$, satisfying the following three conditions:

(S.1) $\sigma(t,\omega)$ is measurable in (t,ω),

(S.2) $\int_\alpha^\beta \sigma(\tau,\omega)^2\,d\tau<\infty$ for almost all ω, and

(S.3) for any t, $\alpha\leq t\leq\beta$, the system $(\sigma(\tau,\omega), \alpha\leq\tau\leq t;$ $g(\tau,\omega)-g(\alpha,\omega), \alpha\leq\tau\leq t)$ is independent of $(g(\tau,\omega)-g(t,\omega),\ t\leq\tau\leq\beta)$. As is easily verified, $S(I_1)$ is conditionally complete; if $\sigma_n\in S(I_1)$ tends to σ_∞ for almost all (t,ω) and if $|\sigma_n|\leq\sigma_o\in S(I_1)$, then $\sigma_\infty\in S(I_1)$.

Theorem 7.1. We can determine, for $\sigma\in S(I_1)$,

(7.1') $\int_\alpha^\beta \sigma(\tau,\omega)dg(\tau,\omega)$ or $\int_{I_1}(\tau,\omega)dg(\tau,\omega)$ or briefly $\int(\sigma,\omega)$ in one and only one way so that it may satisfy (G.1) and (G.2). Furthermore it satisfies (G.3), (G.4), (G.5) and (G.6).

(G.1) When $\sigma(t,\omega)$ is a uniformly stepwise function, i.e., when there exist $\alpha=t_o<t_1<\ldots<t_k=\beta$ independent of ω such that $\sigma(t,\omega)=\sigma(t_{\nu-1},\omega), t_{\nu-1}\leq t<t_\nu$, we have

$$\int(\sigma,\omega)=\sum_{\nu=1}^{k}\sigma(t_{\nu-1},\omega)(g(t_\nu,\omega)-g(t_{\nu-1},\omega)).$$

(G.2) If $\sigma_n\in S(I_1)$ tends to σ_∞ for almost all (τ,ω), and if $|\sigma_n|\leq\sigma_o\in S(I_1)$ and further if every B-measurable function $\sigma(t,\omega)$ of $(\sigma_1,\sigma_2,\ldots)$

satisfies (S.3), then $\int(\sigma_n, \omega)$ converges to $\int(\sigma_\infty, \omega)$ in probability.

(G.3) $$\int(c_1\sigma_1 + c_2\sigma_2, \omega) = c_1 \int(\sigma_1, \omega) + c_2 \int(\sigma_2, \omega)$$

if $\sigma_1, \sigma_2, c_1\sigma_1 + c_2\sigma_2 \in S(I_1)$.

(G.4) $$\mathcal{E}((\int_{I_1}(\sigma, \omega))^2) = \int_{I_1}\mathcal{E}(\sigma^2(\tau, \omega))d\tau$$

if the right side is finite.

(G.5) If $\sigma_1(\tau, \omega) = \sigma_2(\tau, \omega)$ for $\tau \in I_1$, $\omega \in \Omega_1$, Ω_1 being a P-measurable set, then $\int(\sigma_1, \omega) = \int(\sigma_2, \omega)$ for almost all (P) $\omega \in \Omega_1$.

(G.6) If $\int_{I_1}\mathcal{E}(\sigma^2(\tau, \omega))d\tau < \infty$, then $\mathcal{E}(\int(\sigma, \omega)) = 0$.

Proof of the existence. In case σ is a uniformly stepwise function we define by (G.1). It is evident that this definition satisfies (G.3), (G.4), (G.5) and (G.6). The condition (S.3) will be used in the proof of (G.4) and (G.5) and (G.6).

In order to define $\int(\sigma, \omega)$ for $\sigma \in S(I_1)$ such that

(7.2) $$\|\sigma\|^2 = \int_{I_1}\mathcal{E}(\sigma^2(\tau, \omega))d\tau < \infty,$$

we shall establish

Lemma 7.1. For any $\sigma \in S(I)$ satisfying (7.2) we can find a sequence of uniformly stepwise functions $\sigma_n \in S(I)$ such that

(7.3) $$\|\sigma_n - \sigma\|^2 = \int_{I_1}\mathcal{E}((\sigma_n(\tau, \omega) - \sigma(\tau, \omega))^2)d\tau$$

may tend to 0.

The proof can be achieved by the method[26] J. L. Doob has used in his research of measurable stochastic processes. By defining $\sigma(\tau, \omega) = 0$ for $\tau \leq \alpha$ or $\tau > \beta$, we may assume that $\sigma(\tau, \omega) \in L_2(R^1 \times \Omega)$, and so, for almost all ω, $\sigma(\tau, \omega) \in L_2(R^1)$.