Itô's theorems on chaos expansions and martingale representations

Shinzo Watanabe

1. Introduction.

So many comments and appreciations have been given on magnificent contributions by Kiyosi Itô (1915–2008) in probability theory. And yet, I would like to repeat again, in the year of *Itô's centenary*, some of remarkable and characteristic features of his work in stochastic analysis.

As he recollected in [**It 8**], he came across with Kolmogorov's book, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, when he was a student of the University of Tokyo around 1935. Kolmogorov founded the probability theory, as we know well, on a rigorous basis of modern analysis such as the theory of measures and integrations. He formulated random phenomena in natural and social worlds as mathematical models called *stochastic processes*. In his formulation, a stochastic process is a collection of *random variables* defined on a *probability space*, depending on time, taking their values in a finite dimensional state space, typically, in a Euclidean space. A stochastic process is determined uniquely by the system of *joint distributions* of random variables at finite different time points and, conversely, it can be constructed from a given consistent family of distributions so that they coincide with joint distributions by Kolmogorov's extension theorem. Thus, a Markov process, which is a particular and most important case of stochastic processes, is determined by and constructed from the system of *transition probabilities*. Kolmogorov, then one of his followers, W. Feller, obtained the fundamental evolution equations governing transition probabilities so that Markov processes can be obtained and studied completely through the *equations of Kolmogorov and Feller*.

In a striking contrast with such an analytical approach to Markov processes, P. Lévy, who is also a great pioneer of modern probability theory, gave another one, which may be called a *sample paths theoretic approach*. Having successfully determined the canonical form of infinitely divisible distributions, known now as the *Lévy-Khinchin canonical form*, he rewrote this result in terms of sample functions of Lévy processes, (additive processes, differential processes, or processes with independent increments). He regarded a stochastic processes as a random variable taking its values in a function space (a space of paths); a sample function is each value taken by random variable. He represented a sample function of a Lévy process as a sum of a sample function of *Wiener process*, (of which N. Wiener, still another great pioneer of the modern probability theory, has successfully constructed in 1923) and a certain random function given by an integral with respect to *Poisson random measures*.

Itô was interested in this Lévy's approach: In studying carefully the work of Lévy which is deep but looking somewhat a little intuitive, he wished to complete it on a rigorous foundation given by Kolmogorov: He succeeded in this work and published it in [**It 1**]. He then applied the same idea to sample functions of Markov processes determined by Kolmogorov-Feller's equation. By introducing an original notion of *stochastic* *integrals* by sample functions of Wiener processes and Poisson random measures, he finally succeeded in writing down the fundamental equation for a sample function of Markov process, (cf. [**It 2**], [**It 3**]). Thus, *Itô's theory of stochastic differential equations* has been completed: in this theory, he established a kind of Newton-Leibniz differential and integral calculus which can analyze sample functions of stochastic processes (cf. e.g. [**It 4**], [**It 7**]). We will not review on Itô's calculus in this note: So many textbooks and monographs have treated on this subject and so many comments and appreciations have been given. In the next section, we review on one topic in Itô's calculus, his theory of chaos expansions, as a main mathematical content of this note.

2. Chaos Expansions.

In the following, all stochastic processes are given on some probability space $(\Omega, \mathcal{F}, P)$ although we usually suppress stochastic parameter $\omega \in \Omega$.

For $n = 1, 2, \ldots, d$, $1 \leq d \leq \infty$, let $w(t) = (w(t))$, $0 \leq t < \infty$ be the d-dimensional Wiener process: $w(t) = (w_1(t), \ldots, w_d(t))$.[1] Set

$$\mathcal{F}^w_{s,t} = \sigma[w(v) - w(u); s \leq u \leq v \leq t].^2$$

Then, obviously, for $0 \leq r \leq s \leq t$, σ-fields $\mathcal{F}^w_{r,s}$ and $\mathcal{F}^w_{s,t}$ are independent and generate the σ-field $\mathcal{F}^w_{r,t}$. We denote this fact as

$$\mathcal{F}^w_{r,s} \otimes \mathcal{F}^w_{s,t} = \mathcal{F}^w_{r,t}. \tag{2.1}$$

The family $\{\mathcal{F}^w_{s,t}\}$ of σ-fields is called a *Wiener noise.*

Let $(X, \mathcal{B}_X, n(dx))$ be a σ-finite measure space and let $p = (p(t), t \in \mathbb{D}_p)$ be a stationary *Poisson point process* on X *with the characteristic measure* $n(dx)$ as defined in Definition 9.1 of [**Ik-Wa**, p. 43]. When we consider a Wiener process and Poisson point process together on a same probability space, we always assume that *they are given independently.* Set, for $0 \leq s \leq t < \infty$,

$$\mathcal{F}^p_{s,t} = \sigma[N_p((u, v] \times E) | s \leq u \leq v \leq t, E \in \mathcal{B}_X]$$

where $N_p(A) = \#\{t \in \mathbb{D}_p | (t, p(t)) \in A\}$ is the counting measure associated with the point process p. Then we have obviously

$$\mathcal{F}^p_{r,s} \otimes \mathcal{F}^p_{s,t} = \mathcal{F}^p_{r,t} \quad \text{for } 0 \leq r \leq s \leq t. \tag{2.2}$$

The family $\{\mathcal{F}^p_{s,t}\}$ of σ -fields is called a *Poisson noise.*

Now we review on Itô's works on chaos expansion for functionals of Wiener and Poisson noises following [**It 5**] and [**It 6**]. Certainly, Itô has been motivated by the work of N. Wiener [**Wi 1**]: Wiener's work was incomplete in the point that chaos integrals

[1] In the case $d = \infty$, this should be written as $w(t) = (w_1(t), \ldots, w_n(t), \ldots.)$, since there is no d-component. This convention is always used below without mentioning.

[2] We set $\mathcal{F}^w_{s,\infty} = \bigvee_{t \geq s} \mathcal{F}^w_{s,t}$.

of different orders are not necessarily orthogonal and Itô corrected this by introducing the notion of *multiple Wiener integrals*. Multiple Wiener integrals can be equivalently defined, as Itô himself remarked in Section 5 of [**It 5**], by *successive Itô's stochastic integrals*. We follow here this second approach of Itô; it has a merit, for example, that the martingale representation theorem can be almost immediately obtained from this. It should be remarked that Wiener's monograph [**Wi 2**] corrected his incomplete chaos integrals and discussed many interesting and important applications.

Let a Wiener process w and a Poisson point process p be given. Let $\mathcal{F}^{(w,p)}_{s,t}$ be the σ-field generated by mutually independent σ-fields $\mathcal{F}^{w}_{s,t}$ and $\mathcal{F}^{p}_{s,t}$. Then we have

$$\mathcal{F}^{(w,p)}_{r,s} \otimes \mathcal{F}^{(w,p)}_{s,t} = \mathcal{F}^{(w,p)}_{r,t}, \ r \le s \le t.^3 \tag{2.3}$$

For a sub σ-field $\mathcal{G}$ of $\mathcal{F}$, we denote by $L^2(\mathcal{G})$ the real L^2-space formed of all real, $\mathcal{G}$-measurable and square-integrable random variables. The chaos expansion theorems are concerned with the orthogonal decomposition of the space $L^2(\mathcal{F}^{(w,p)}_{s,t})$.

Set $\hat{X} = X \uplus \{1, 2, \dots, d\}^4$, $\mathcal{B}_{\hat{X}} = \{E \subset \hat{X} | E \cap X \in \mathcal{B}_X\}$ and $\hat{n}(dx)$ be a σ-finite measure on $(\hat{X}, \mathcal{B}_{\hat{X}})$ defined by

$$\hat{n}(E) = n(E \cap X) + \sum_{i=1}^{d} 1_{E \cap \{1,2,\dots,d\}}(i).$$

Define the compensated counting measure $\tilde{N}_p$ by

$$\tilde{N}_p(dtdx) = N_p(dtdx) - dtn(dx).$$

For each fixed $0 \le s < \infty$, define a filtration $\mathbb{F}^{(s)}$ by

$$\mathbb{F}^{(s)} = \big\{\mathcal{F}^{(w,p)}_{s,t}; s \le t\big\}.$$

Then, for an $\mathbb{F}^{(s)}$-predictable $f(\tau)$, $\tau \ge s$, such that $E\big[\int_s^t |f(\tau)|^2 d\tau\big] < \infty$, Itô's stochastic integral $\int_s^t f(s) dw_i(s)$ is defined for $i = 1, 2, \dots, d$ and, for $\mathbb{F}^{(s)}$-predictable process $f(\tau, x)$ on $[s, \infty) \times X$ in the sense of Definition 3.3 of [**Ik-Wa**, p. 61] such that $E\big[\int_s^t d\tau \int_X |f(\tau, x)|^2 n(dx)\big] < \infty$, stochastic integral

$$\int_s^t \int_X f(\tau, x) \tilde{N}_p(d\tau, dx)$$

is defined as in p. 63 of [**Ik-Wa**]. Hence, if $f(\tau, x)$ on $[s, \infty) \times \hat{X}$ is $\mathbb{F}^{(s)}$-predictable and

[3] For convenience, we allow the case $w \equiv 0$ or $p \equiv 0$ so that $\mathcal{F}^{(w,0)}_{s,t}$ coincides with $\mathcal{F}^{w}_{s,t}$ and $\mathcal{F}^{(0,p)}_{s,t}$ with $\mathcal{F}^{p}_{s,t}$.

[4] $\biguplus$ is the disjoint union (i.e., a sum).

$$E\Big[\int_s^t d\tau \int_{\hat{X}} |f(\tau,x)|^2 \hat{n}(dx)\Big] < \infty,$$

we can define the stochastic integral

$$\int_s^t \int_{\hat{X}} f(\tau,x) M(d\tau,dx) := \sum_{i=1}^d \int_s^t f(\tau,i)dw_i(\tau) + \int_s^t \int_X f(\tau,x)\tilde{N}(d\tau,dx) \tag{2.4}$$

and it holds that

$$\begin{aligned} E\Big[\Big(\int_s^t \int_{\hat{X}} f(\tau,x)M(d\tau,dx)\Big)^2\Big] &= E\Big[\int_s^t d\tau \int_{\hat{X}} |f(\tau,x)|^2\, \hat{n}(dx)\Big] \\ &= \sum_{i=1}^d E\Big[\int_s^t |f(\tau,i)|^2\, d\tau\Big] + E\Big[\int_s^t d\tau \int_X |f(\tau,x)|^2\, n(dx)\Big]. \end{aligned} \tag{2.5}$$

Let, for each $n = 1, 2, \dots$ and $0 \le s < t \le \infty$,

$$\begin{aligned} \Delta_{s,t}^{(n)} = \{&(t_1, t_2, \dots, t_{n-1}, t_n; x_1, \dots, x_{n-1}, x_n) \\ &\mid s < t_1 < t_2 < \cdots < t_{n-1} < t_n < t,\ x_i \in \hat{X}, i = 1, \dots, n\} \end{aligned}$$

and $L^2(\Delta_{s,t}^{(n)}; dt_1, dt_2, \dots, dt_{n-1}, dt_n, \hat{n}(dx_1)\cdots\hat{n}(dx_{n-1})\hat{n}(dx_n))$[5] be the real L^2-space formed of real, square-integrable, measurable functions $f = (f(t_1, t_2, \dots, t_{n-1}, t_n; x_1, \dots, x_{n-1}, x_n))$ on $\Delta_{s,t}^{(n)}$ with the norm

$$\begin{aligned} \|f\|_{L^2(\Delta_{s,t}^{(n)})} = \Big(\int_s^t dt_n \int_{\hat{X}} \hat{n}(dx_n)\Big[\int_s^{t_n} dt_{n-1} \int_{\hat{X}} \hat{n}(dx_{n-1}) \\ \Big[\cdots\Big[\int_s^{t_2} dt_1 \int_{\hat{X}} \hat{n}(dx_1) f(t_1, \dots, t_{n-1}, t_n; x_1, \dots, x_{n-1}, x_n)^2\Big]\cdots\Big]\Big]\Big)^{1/2}. \end{aligned}$$

Then, for $f \in L^2(\Delta_{s,t}^{(n)})$, we can define the iterated stochastic integral $I^{(n)}(f)(s,t)$ by

$$\begin{aligned} &I^{(n)}(f)(s,t) \\ &:= \int_s^t \int_{\hat{X}} M(dt_n, dx_n)\Big[\int_s^{t_n-} \int_{\hat{X}} M(dt_{n-1}, dx_{n-1}) \\ &\quad \cdots\Big[\int_s^{t_2-} \int_{\hat{X}} f(t_1, t_2, \dots, t_{n-1}, t_n; x_1, x_2, \dots, x_{n-1}, x_n) M(dt_1, dx_1)\Big]\cdots\Big] \end{aligned} \tag{2.6}$$

[5] This is often denoted by $L^2(\Delta_{s,t}^{(n)})$ to simplify the notation.

by noting that, in each step, integrands are $\mathbb{F}^{(s)}$-predictable processes. Then we have $I^{(n)}(f)(s,t) \in L^2(\mathcal{F}_{s,t}^{(w,p)})$ with

$$\|I^{(n)}(f)(s,t)\|_{L^2(\mathcal{F}_{s,t}^{(w,p)})} = \|f\|_{L^2(\Delta_{s,t}^{(n)})}. \tag{2.7}$$

It holds that, if $n \neq m, f \in L^2(\Delta_{s,t}^{(n)})$ and $g \in L^2(\Delta_{s,t}^{(m)})$, then $I^{(n)}(f)$ and $I^{(m)}(g)$ are orthogonal in $L^2(\mathcal{F}_{s,t}^{(w,p)})$.

The following result was essentially obtained[6], for the case $p = 0$ (i.e. the case of Wiener noise) in [**It 5**] and, for the general case in [**It 6**].

THEOREM 2.1. *If $f \in L^2(\mathcal{F}_{s,t}^{(w,p)})$, there exists a sequence $f^{(n)} = (f^{(n)}(t_1, \ldots, t_n, x_1, \ldots, x_n)) \in L^2(\triangle_{s,t}^{(n)})$, $n = 1, 2, \ldots$, satisfying*

$$\sum_{n=1}^{\infty} \|f^{(n)}\|^2_{L^2(\triangle_{s,t}^{(n)})} < \infty \tag{2.8}$$

such that the following expansion holds:

$$f = E(f) + \sum_{n=1}^{\infty} I^n(f^{(n)})(s,t). \tag{2.9}$$

By the orthogonality of $I^{(n)}$ for different n, the expansion (2.9) is, in fact, an orthogonal expansion in $L^2(\mathcal{F}_{s,t}^{(w,p)})$.

An important consequence is that the following representation theorem for $\mathbb{F}^{(s)}$-martingale can be obtained:

THEOREM 2.2. *Every square-integrable $\mathbb{F}^{(s)} = \{\mathcal{F}_{s,t}^{(w,p)} : t \geq s\}$-martingale $M = (M_t)_{t \geq s}$ can be represented in the form*

$$\begin{aligned} M_t &= E[M_s] + \int_s^t \int_{\hat{X}} \Phi(\tau, x) M(d\tau, dx) \\ &= E[M_s] + \sum_{i=1}^{d} \int_s^t \Phi_i(\tau) dw_i(\tau) + \int_s^t \int_X \Psi(\tau, x) \tilde{N}_p(d\tau, dx) \end{aligned} \tag{2.10}$$

by the uniquely determined predictable integrands $\Phi = (\Phi_i, \Psi)$.

This result in the case of $p = 0$, i.e. the case of Wiener process, is now well-known as *Itô's representation theorem.*

The *Clark-Ocone* formula, which is concerned with the integrand $\Phi_i(\tau)$ in (2.10), can also be obtained from the expansion (2.9), Cf. [**Wa**].

---

[6] The only difference is that the Poisson point process p is a little bit restricted in [**It 6**]; in our case it is completely general.

Following Tsirelson (cf. [**T 1**], [**T 2**], [**T 3**]) we call a family $\{\mathcal{F}_{s,t}\}_{0\le s\le t<\infty}$ of sub σ-fields of $\mathcal{F}$ a *noise* if it satisfies $\mathcal{F}_{t,t}=\{\emptyset,\Omega\}$ module P-null set, $\mathcal{F}_{r,t}=\mathcal{F}_{r,s}\otimes\mathcal{F}_{s,t}$ for $r\le s\le t$, and for $h>0$, $\mathcal{F}_{s,t}$ and $\mathcal{F}_{s+h,t+h}$ are isomorphic under a measure preserving transformation T_h on $(\Omega,\mathcal{F})$. Thus, $\{\mathcal{F}^{(w,p)}_{s,t}\}$ is a typical example of noise. Tsirelson defined a noise *linear* (*linearizable*, *classical* or *white*) if it is given as $\{\mathcal{F}^{(w,p)}_{s,t}\}$ by some Wiener process w and Poisson point process p. Thus, if $l=(l(t))_{t\ge0}$ is a time homogeneous Lévy process on $\mathbb{R}^d$, i.e. a right-continuous process with stationary independent increments, then

$$\mathcal{F}^l_{s,t}=\sigma[l(v)-l(u); s\le u\le v\le t]$$

defines a noise $\{\mathcal{F}^{(l)}_{s,t}\}$. By Lévy–Itô's decomposition theorem we see that $\mathcal{F}^{(l)}_{s,t}=\mathcal{F}^{(w,p)}_{s,t}$ by a Wiener process and a Poisson point proccess on $\mathbb{R}^d\setminus\{0\}$ so that $\mathcal{F}^{(l)}_{s,t}$ is a linear noise.

Given a noise, Tsirelson introduced an important notion called the *spectrum of noise*. Let $\mathcal{C}_{[0,\infty)}$ be the space of all compact subsets in $[0,\infty)$ endowed with the Hausdorff metric and $\mathcal{C}^{\text{finite}}$ be its subspace formed of all finite subsets (including the empty set). By an elementary set B, we mean a finite disjoint union of compact intervals in $[0,\infty)$: $B=\bigcup_{i=1}^m[u_i,v_i]$. Set $\mathcal{F}_B=\mathcal{F}_{u_1,v_1}\vee\mathcal{F}_{u_2,v_2}\vee\cdots\vee\mathcal{F}_{u_m,v_m}$.

Let $\mathbb{N}=\{\mathcal{F}_{s,t}\}$ be a given noise. Then, for every $f\in L^2(\mathcal{F}_{0,\infty})$ with $\|f\|_{L^2}=1$, there exists a unique Borel probability μ_f on $\mathcal{C}_{[0,\infty)}$ with the property

$$\mu_f\big(\{F\in\mathcal{C}_{[0,\infty)}; F\subset B\}\big)=E[E(f|\mathcal{F}_B)^2].$$

μ_f is called the *spectral measure* of the noise $\{\mathcal{F}_{s,t}\}$ with respect to f.

In the case of the noise $\{\mathcal{F}^{(w,p)}_{s,t}\}$, μ_f can be obtained by the chaos expansion formula (2.9) for $f\in L^2(\mathcal{F}_{0,\infty})$ with $\|f\|_{L^2(\mathcal{F}_{0,\infty})}=1$. In this case, μ_f has its full measure on $\mathcal{C}^{\text{finite}}$ such that $\mu_f\{\emptyset\}=E[f]^2$, and letting $S_m=\{a_1,a_2,\dots,a_m\}\in\mathcal{C}^{\text{finite}}$, $0\le a_1<\cdots<a_m<\infty$,

$$\begin{aligned}&\mu_f\{S_m\in dt_1dt_2\dots dt_m\}\\&\quad=\left[\int_{\hat{X}}\cdots\int_{\hat{X}}|f^{(m)}(t_1,\dots,t_m;x_1,\dots,x_m)|^2\hat{n}(dx_1)\dots\hat{n}(dx_m)\right]dt_1\dots dt_m,\\&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad m=1,2,\dots.\end{aligned}$$

Tsirelson obtained the following important fact: A noise $\{\mathcal{F}_{s,t}\}$ is linear if and only if the spectral measure μ_f associated with $f\in L^2(\mathcal{F}_{0,\infty})$ such that $E[f^2]=1$ has its full measure on $\mathcal{C}^{\text{finite}}$ for every f, cf [**T 2**] or [**T 3**].

References

[Ik-Wa] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, 2nd. Edition, North-Holland/Kodansha, 1989.

[It 1] K. Itô, On stochastic processes (infinitely divisible laws of probability), Japan Jour. Math.,

XVIII (1942), 261–301.
[It 2] K. Itô, Differential equations determining a Markov process (Original Japanese: Zenkoku Sizyo Sugaku Danwakai-si), Pan-Japan Math. Coll. No. 1077 (1942).
[It 3] K. Itô, On stochastic differential equations, Mem. Amer. Math. Soc., **4** (1950), 1–51.
[It 4] K. Itô, On a formula concerning stochastic differentials, Nagoya Math. Jour., **3** (1951), 55–65.
[It 5] K. Itô, Multiple Wiener integral, Jour. Math. Soc. Japan, **3** (1951), 157–169.
[It 6] K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, Trans. Amer. Math. Soc., **81** (1956), 253–263.
[It 7] K. Itô, Stochastic differentials, Appl. Math. and Opt., **1** (1974), 374–381.
[It 8] K. Itô, A note of thanks by K. Itô, In: Stochastic Analysis and Related Topics in Kyoto (in honor of Kiyosi Itô), Adv. Studies in Pure Math., **41** Math. Soc. Japan (2004) in the page after Preface.
[T 1] B. Tsirelson, Within and beyond the reach of Brownian innovation, "Documenta Mathematica" Extra Volume ICM 1998, **III**, 311–326.
[T 2] B. Tsirelson, Unitary Brownian motions are linearizable, **PR**/9806112 (1998).
[T 3] B. Tsirelson, Nonclassical stochastic flows and continuous products, Probability Surveys, **1** (2004), 173–298.
[Wa] S. Watanabe, Martingale representation theorem and chaos expansion, In: Stochastic Processes and Applications to Mathematical Finance, World Scientific, 2006, 195–217.
[Wi 1] N. Wiener, The homogeneous chaos, Amer. Jour. Math., **60** (1938), 897–936.
[Wi 2] N. Wiener, Nonlinear Problems in Random Theory, The MIT Press, Cambridge Mass, and Wiley, New York, 1958.