Communications on Stochastic Analysis

Volume 8 | Number 1 Article 8

3-1-2014

The Itô calculus and white noise theory: a brief survey toward general stochastic integration

Hui-Hsiung Kuo

Communications on Stochastic Analysis
Vol. 8, No. 1 (2014) 111-139

Serials Publications
www.serialspublications.com

THE ITÔ CALCULUS AND WHITE NOISE THEORY: A BRIEF SURVEY TOWARD GENERAL STOCHASTIC INTEGRATION

HUI-HSIUNG KUO*

Abstract. We give a brief survey on the Itô calculus and white noise theory with the aim to extend the Itô theory of stochastic integration to stochastic processes which may not be adapted. The Hitsuda–Skorokhod integral by white noise methods provides only a partial extension of the Itô integral. A new class of stochastic processes, called instantly independent, is introduced and serves as a counterpart of adapted stochastic processes in the Itô theory. Then a new stochastic integral is defined for linear combinations of products of adapted and instantly independent stochastic processes. We explain recent results on this new theory and mention several open problems for further investigation and study.

1. A Simple Anticipating Stochastic Differential Equation

Let $B(t)$ be a Brownian motion and let $\{\mathcal{F}_t;\, t \geq 0\}$ be the associated filtration, i.e., $\mathcal{F}_t = \sigma\{B(s), 0 \leq s \leq t\}$. For $a \geq 0$, consider the following stochastic differential equation

$$dX_t = X_t\, dB(t), \quad X_a = \xi, \quad t \geq a, \tag{1.1}$$

where ξ is $\mathcal{F}_a$-measurable. It is well known that the solution of this equation is given by

$$X_t = \xi e^{B(t)-B(a)-\frac{1}{2}(t-a)}, \quad t \geq a. \tag{1.2}$$

In particular, when $a = 0$ and $\xi = x \in \mathbb{R}$, we have the solution

$$X_t = xe^{B(t)-\frac{1}{2}t}, \quad t \geq 0. \tag{1.3}$$

The stochastic process $X_t, t \geq 0$, in Equation (1.3) is a martingale and a Markov process. In fact, the martingale property and the Markov property are the two guiding properties for which Itô developed the theory of stochastic integration in the very beginning in 1942 [6].

Received 2014-5-30; Communicated by the editors.
2010 *Mathematics Subject Classification.* Primary 60H05, 60H40; Secondary 60H07, 60H20.
Key words and phrases. Brownian motion, Wiener integral, isometry, adapted stochastic process, Itô integral, quadratic variation of Brownian motion, martingale, Itô's formula, Markov process, multiple Wiener-Itô integral, white noise, white noise space, generalized white noise functions, S-transform, characterization theorem, white noise differentiation, adjoint operator, white noise integral, Hitsuda–Skorokhod integral, instantly independent stochastic process, new stochastic integral, near-martingale, Girsanov theorem, exponential process.

* This article is based on two lectures given at the Summer School and a lecture delivered at the International Conference during the Third Buea International Conference on Mathematical Sciences, April 22–May 3, 2013, Buea, Cameroon.

Now, consider the modified stochastic differential equation

$$dY_t = Y_t\,dB(t), \quad Y_0 = B(1), \quad t \ge 0, \tag{1.4}$$

where the initial condition $Y_0 = B(1)$ is anticipating. Clearly, this equation is not within the Itô theory of stochastic integration. But then what is the solution of this simple stochastic differential equation in Equation (1.4)?

We may try to use the iteration method to derive the solution. First consider the case when $0 \le t \le 1$. Let $Y_t^{(0)} = B(1)$ and define

$$\begin{aligned} Y_t^{(1)} &= B(1) + \int_0^t Y_s^{(0)}\,dB(s) \\ &= B(1) + \int_0^t B(1)\,dB(s). \end{aligned}$$

Observe that the integral $\int_0^t B(1)\,dB(s)$ is not an Itô integral since the integrand $B(1)$ is not adapted with respect to the filtration $\{\mathcal{F}_t;\ t \ge 0\}$. In §4 we will introduce an extension of Itô integral and show in Equation (4.4) the following stochastic integral

$$\int_0^t B(1)\,dB(s) = B(1)B(t) - t, \quad 0 \le t \le 1.$$

Therefore, we have

$$Y_t^{(1)} = B(1) + B(1)B(t) - t, \quad 0 \le t \le 1.$$

We go one more step to define $Y_t^{(2)}$,

$$\begin{aligned} Y_t^{(2)} &= B(1) + \int_0^t Y_s^{(1)}\,dB(s) \\ &= B(1) + \int_0^t \big(B(1) + B(1)B(s) - s\big)\,dB(s). \end{aligned} \tag{1.5}$$

Again, for our extension of the Itô integral in Section 4, by Equation (4.5) we have the stochastic integral

$$\int_0^t B(1)B(s)\,dB(s) = \frac{1}{2}B(1)\big(B(t)^2 - t\big) - \int_0^t B(s)\,ds. \tag{1.6}$$

Then we can put Equation (1.6) into Equation (1.5) to get

$$Y_t^{(2)} = B(1) + \big(B(1)B(t) - t\big) + \frac{1}{2}B(1)\big(B(t)^2 - t\big) - tB(t), \quad 0 \le t \le 1.$$

In general, we can derive that

$$Y_t^{(n)} = B(1)\sum_{k=0}^{n} \frac{1}{k!}H_k\big(B(t);t\big) - t\sum_{k=0}^{n-1} \frac{1}{k!}H_k\big(B(t);t\big), \tag{1.7}$$

where $H_n(x;\sigma^2)$ is the Hermite polynomial of degree n with parameter σ^2, i.e.,

$$H_n(x;\sigma^2) = (-\sigma^2)^n e^{x^2/2\sigma^2} D_x^n e^{-x^2/2\sigma^2}.$$

Note that the generating function of Hermite polynomials is given by

$$e^{tx-\frac{1}{2}\sigma^2 t^2} = \sum_{k=0}^{\infty} \frac{t^k}{k!} H_k(x;\sigma^2).$$

Hence by letting $n \to \infty$ in Equation (1.7), we see that

$$Y_t = \big(B(1) - t\big)e^{B(t)-\frac{1}{2}t}, \quad 0 \le t \le 1. \tag{1.8}$$

Next consider the case when $t \ge 1$, i.e., the stochastic differential equation

$$dY_t = Y_t\, dB(t), \quad Y_1 = \big(B(1) - 1\big)e^{B(1)-\frac{1}{2}}, \quad t \ge 1. \tag{1.9}$$

Obviously, this stochastic differential equation is within the Itô theory of stochastic integration and by Equation (1.2) the solution is given by

$$\begin{aligned} Y_t &= \big(B(1) - 1\big)e^{B(1)-\frac{1}{2}} e^{B(t)-B(1)-\frac{1}{2}(t-1)} \\ &= \big(B(1) - 1\big)e^{B(t)-\frac{1}{2}t}, \quad t \ge 1. \end{aligned} \tag{1.10}$$

Putting Equations (1.8) and (1.10) together, we get the solution of the anticipative stochastic differential equation in (1.4):

$$Y_t = \begin{cases} \big(B(1) - t\big)e^{B(t)-\frac{1}{2}t}, & \text{if } 0 \le t \le 1, \\ \big(B(1) - 1\big)e^{B(t)-\frac{1}{2}t}, & \text{if } t > 1. \end{cases}$$

Observe that the stochastic process Y_t for $0 \le t \le 1$ is not a martingale, nor a Markov process. However, the stochastic process Y_t for $t \ge 1$ is a martingale and is also a Markov process. Thus we raise the question: What kind of new properties does the stochastic process Y_t, $0 \le t \le 1$, have?

2. The Itô Calculus

In this section we briefly review the Itô theory of stochastic integration.

2.1. Original motivation. Here is the scenario at the time when Itô introduced his theory of stochastic integration in 1942 [6].

(1) Suppose X_t is a stochastic process. Then we have a family of marginal distributions of X_t. Conversely, given a family of distributions satisfying the consistency condition, we can apply the Kolmogorov extension theorem to obtain a stochastic process X_t such that the given distributions are the marginal distributions of X_t. In particular, if X_t is a stationary diffusion process with infinitesimal generator $\mathcal{A}$, then we have a family of transition probabilities $\{p_t(x,\cdot)\}$ satisfying the Chapman-Kolmogorov equation.

(2) Suppose $\{p_t(x,\cdot)\}$ is a family of transition probabilities satisfying the Chapman-Kolmogorov equation. Then we can define a C_0-contraction semigroup $\{P_t\}$. Conversely, given a semigroup $\{P_t\}$, we can use the Riesz representation theorem to obtain a family of transition probabilities satisfying the Chapman-Kolmogorov equation such that their associated semigroup is the given $\{P_t\}$.

(3) At the same time K. Yosida was working on constructing a semi-group $\{P_t\}$ from an infinitesimal generator $\mathcal{A}$, which later became the well-known Hille-Yosida theorem. (*From a private conversation with K. Itô: It was his mentor K. Yosida who asked him to find a probabilistic method to construct diffusion processes from an infinitesimal generator.*)

Now, with a given infinitesimal operator $\mathcal{A}$, how can we construct a stochastic process X_t such that the associated infinitesimal generator is $\mathcal{A}$? In view of the above description (1) (2) (3), we will need to apply the Hille-Yosida theorem to find a semigroup $\{P_t\}$. Then use the Riesz representation theorem to find transition probabilities $\{p_t(x, \cdot)\}$. And finally use the Kolmogorov extension theorem to construct a stochastic process X_t whose infinitesimal generator is $\mathcal{A}$. Alternatively, we can avoid using the Hille-Yosida theorem by solving the Kolmogorov forward or backward equation to find the transition probabilities $\{p_t(x, \cdot)\}$. And then apply the Kolmogorov extension theorem to construct such a stochastic process X_t. For detail, see Section §10.9 in the book [14].

The original motivation for K. Itô to introduce the Itô integral is to directly construct a stochastic process from a given infinitesimal generator $\mathcal{A}$ without going through the various steps as described above. By thinking of the displacement ΔX_t of a stochastic process X_t as depending on the displacements $\Delta B(t)$ of a Brownian motion and Δt of the time parameter, K. Itô came up with the following stochastic differential equation

$$dX_t = f(t, X_t)\, dB(t) + g(t, X_t)\, dt, \quad X_0 = x,$$

which is interpreted as meaning the stochastic integral equation

$$X_t = x + \int_0^t f(s, X_s)\, dB(s) + \int_0^t g(s, X_s)\, ds, \tag{2.1}$$

since almost all Brownian motion paths are nowhere differentiable.

For a given infinitesimal generator $\mathcal{A}$, we have the functions f and g. Then we can solve Equation (2.1) to get a stochastic process X_t in just one step. However, here are crucial points:

(1) In Equation (2.1), what is the definition of the first integral with respect to a Brownian motion $B(t)$?
(2) How do we know that the solution X_t in Equation (2.1) is a Markov process and its infinitesimal generator is the given operator $\mathcal{A}$?

All of the above discussions led K. Itô [6] to introduce the following stochastic integral in 1942:

$$\int_a^b f(t)\, dB(t), \tag{2.2}$$

where the integrand $f(t)$ is a stochastic process. Of course, we need to impose conditions on $f(t)$. We quickly explain this stochastic integral. For detail, see the book [14].

2.2. Brownian motion. First recall that a *Brownian motion* is a stochastic process $B(t, \omega)$, $t \geq 0, \omega \in \Omega$, satisfying the following conditions:

1. $P\{B(0, \cdot) = 0\} = 1$,

2. $B(s) - B(t)$ is normal with mean 0 and variance $s - t$ for any $0 \le t < s$,
3. $B(t)$ has independent increments,
4. $P\{\omega; B(\cdot, \omega) \text{ is continuous}\} = 1$.

By the second condition, we see that $E\big[(B(s) - B(t))^2\big] = |s - t|$, which can be informally interpreted as $|B(s) - B(t)| \approx \sqrt{|s - t|}$. Hence we see that Brownian motion paths are nowhere differentiable. In fact, we have the following well-known theorem on the quadratic variation of a Brownian motion. This theorem is a fundamental property in the Itô theory of stochastic integration.

Theorem 2.1. *Let $B(t)$ be a Brownian motion and let $\Delta_n = \{a = t_0, t_1, t_2, \ldots, t_n = b\}$ be a partition of a finite interval $[a, b]$. Then*

$$\sum_{i=1}^{n} \big(B(t_i) - B(t_{i-1})\big)^2 \longrightarrow b - a$$

in $L^2(\Omega)$ as $\|\Delta_n\| = \max_{1 \le i \le n}(t_i - t_{i-1})$ tends to 0.

2.3. Wiener integral. Nobert Wiener was the first mathematician to give a rigorous construction of a Brownian motion $B(t)$. In his analysis of Brownian functionals, Wiener introduced the homogeneous chaoses of degree $n \ge 0$. In particular, for a function $f \in L^2[a, b]$, he associated a homogeneous chaos $\tilde{f}$ of degree 1 which is a Gaussian random variable with mean 0 and variance $\|f\|^2 = \int_a^b |f(t)|^2 \, dt$. The homogeneous chaos $\tilde{f}$ can be interpreted as the *Wiener integral* of f:

$$\int_a^b f(t) \, dB(t),$$

which is defined in the following two steps:

1. For a step function $f(t)$ given by $f = \sum_{i=1}^{n} a_i 1_{[t_{i-1}, t_i)}$, define

$$I(f) = \sum_{i=1}^{n} a_i \big(B(t_i) - B(t_{i-1})\big).$$

Note that $I(f)$ is well defined and has normal distribution with mean 0 and variance $\|f\|^2 = \int_a^b |f(t)|^2 \, dt$.

2. For $f \in L^2[a, b]$, choose a sequence $\{f_n\}$ of step functions such that $f_n \to f$ in $L^2[a, b]$. It is easy to see that by Step 1, the sequence $\{I(f_n)\}$ is Cauchy in $L^2(\Omega)$. Define

$$\int_a^b f(t) \, dB(t) = \lim_{n \to \infty} I(f_n), \quad \text{in } L^2(\Omega).$$

It is easy to see that the integral $\int_a^b f(t) \, dB(t)$ is well defined.

Theorem 2.2. *Let $f \in L^2[a, b]$. Then the Wiener integral $I(f) = \int_a^b f(t) \, dB(t)$ is a Gaussian random variable with mean 0 and variance $\|f\|^2 = \int_a^b |f(t)|^2 \, dt$.*

It follows from this theorem that for any $f, g \in L^2[a, b]$, we have:

$$E\big(I(f)I(g)\big) = \int_a^b f(t)g(t) \, dt. \tag{2.3}$$

Moreover, by this theorem, the mapping $f \mapsto \int_a^b f(t)\,dB(t)$ is an isometry from $L^2[a,b]$ into $L^2(\Omega)$ and the range consists of Gaussian random variables. However, we should point out that a Brownian functional can be Gaussian and yet not a Wiener integral.

2.4. Itô integral. Consider Equation (1.1) with $a = 0$ and $\xi = x$, namely,

$$dX_t = X_t\,dB(t), \quad X_0 = x, \quad t \geq 0.$$

In attempting to find the solution by the iteration method, we get $X_t^{(0)} = x$ and

$$X_t^{(1)} = x + \int_0^t x\,dB(s) = x(1 + B(t)),$$

$$X_t^{(2)} = x + \int_0^t x\big(1 + B(s)\big)\,dB(s) = x(1 + B(t)) + x\int_0^t B(s)\,dB(s). \tag{2.4}$$

Observe that the stochastic integral $\int_0^t B(s)\,dB(s)$ in Equation (2.4) is not a Wiener integral since the integrand is not a deterministic function.

Here is what K. Itô tried in the beginning of developing his theory of stochastic integration. (*He told me in a private lecture at his home in Ithaca, New York in 1969 when I was a graduate student at Cornell University.*) For a partition $\Delta_n = \{0 = s_0, s_1, \ldots, s_{n-1}, s_n = t\}$ of the interval $[0,t]$, define the following Riemann sums by taking the evaluation points to be the left endpoint and right endpoint of each subinterval:

$$L_n = \sum_{i=1}^n B(s_{i-1})\big(B(s_i) - B(s_{i-1})\big), \quad R_n = \sum_{i=1}^n B(s_i)\big(B(s_i) - B(s_{i-1})\big).$$

It is easy to see that

$$R_n + L_n = B(t)^2,$$

$$R_n - L_n = \sum_{i=1}^n \big(B(s_i) - B(s_{i-1})\big)^2.$$

Therefore, we can solve the equations and then apply Theorem 2.1 to get

$$L_n = \frac{1}{2}\Big(B(t)^2 - \sum_{i=1}^n \big(B(s_i) - B(s_{i-1})\big)^2\Big) \longrightarrow \frac{1}{2}\big(B(t)^2 - t\big), \tag{2.5}$$

$$R_n = \frac{1}{2}\Big(B(t)^2 + \sum_{i=1}^n \big(B(s_i) - B(s_{i-1})\big)^2\Big) \longrightarrow \frac{1}{2}\big(B(t)^2 + t\big) \tag{2.6}$$

in $L^2(\Omega)$ as $\|\Delta_n\| \to 0$. Note that the stochastic process $B(t)^2 - t$ in Equation (2.5) is a martingale, while the stochastic process $B(t)^2 + t$ in Equation (2.6) is not a martingale. Hence in order to keep the martingale property of a Brownian motion $B(t)$ it is more desirable to define the stochastic integral $\int_0^t B(s)\,dB(s)$ so that we have the following equality:

$$\int_0^t B(s)\,dB(s) = \frac{1}{2}\big(B(t)^2 - t\big),$$

namely, the evaluation points are the left endpoints of subintervals.

From the above discussion in attempting to define $\int_0^t B(s)\,dB(s)$, we see that the evaluation of an integrand at the left endpoint of each subinterval yields the martingale property of the integrated stochastic process. In fact, it also yields the Markov property of the solution X_t of the stochastic differential equation in Equation (2.1).

Now, we explain the Itô integral $\int_a^b f(t)\,dB(t)$. Let $B(t)$ be a Brownian motion and $\{\mathcal{F}_t; a \le t \le b\}$ a filtration such that

1. $B(t)$ is $\mathcal{F}_t$-measurable for each t,

2. $B(t) - B(s)$ and $\mathcal{F}_s$ are independent for any $s \le t$.

If $\{\mathcal{F}_t\}$ is not specified, it is understood that $\mathcal{F}_t = \sigma\{B(s); 0 \le s \le t\}$.

Notation. Let $L^2_{\mathrm{ad}}([a,b]\times\Omega)$ denote the Hilbert space of all stochastic processes $f(t,\omega)$ satisfying the conditions:

(1) $f(t)$ is *adapted* to $\{\mathcal{F}_t\}$, i.e., $f(t)$ is $\mathcal{F}_t$-measurable for each $a \le t \le b$,

(2) $\int_a^b E\big(|f(t)|^2\big)\,dt < \infty$.

In the beginning of developing stochastic integration [6, 7], K. Itô defined the stochastic integral

$$\int_a^b f(t)\,dB(t) \tag{2.7}$$

for stochastic processes $f(t)$ in the space $L^2_{\mathrm{ad}}([a,b]\times\Omega)$. The procedure to define this stochastic integral follows the two steps as in the case of Wiener integral. However, for the case when $f(t)$ is a stochastic process, it is much more complicated in the second step because of the requirement of adaptedness.

See Chapter 4 of the book [14] for the definition of the Itô integral $\int_a^b f(t)\,dB(t)$ when $f \in L^2_{\mathrm{ad}}([a,b]\times\Omega)$ and the following Theorems 2.3 to 2.6.

Theorem 2.3. *Let $f \in L^2_{\mathrm{ad}}([a,b]\times\Omega)$ and suppose $E\big(f(t)f(s)\big)$ is a continuous function of t and s. Then*

$$\int_a^b f(t)\,dB(t) = \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n f(t_{i-1})\big(B(t_i) - B(t_{i-1})\big), \quad \text{in } L^2(\Omega), \tag{2.8}$$

where Δ_n's are partitions of $[a,b]$.

In particular, if $f \in L^2_{\mathrm{ad}}([a,b]\times\Omega)$ is a continuous stochastic process, then Equation (2.8) holds.

Theorem 2.4. *Let $f \in L^2_{\mathrm{ad}}([a,b]\times\Omega)$. Then we have the equalities:*

$$E\int_a^b f(t)\,dB(t) = 0,$$

$$E\bigg(\Big|\int_a^b f(t)\,dB(t)\Big|^2\bigg) = \int_a^b E\big(|f(t)|^2\big)\,dt. \tag{2.9}$$

By Equation (2.9) we see that the mapping $f \mapsto \int_a^b f(t)\,dB(t)$ is an isometry from $L^2_{\mathrm{ad}}([a,b]\times\Omega)$ into $L^2(\Omega)$.

Theorem 2.5. *Let $f \in L^2_{\text{ad}}([a,b] \times \Omega)$. Then the stochastic process*

$$X_t = \int_a^t f(s)\,dB(s), \quad a \le t \le b,$$

is a continuous stochastic process, i.e., almost all of its sample paths are continuous functions on the interval $[a,b]$.

Actually, in Theorem 2.5 we need to take a continuous version, namely, there exists a continuous stochastic process $\widetilde{X}_t$, $a \le t \le b$, such that $P\{X_t = \widetilde{X}_t\} = 1$ for each $t \in [a,b]$.

Theorem 2.6. *Let $f \in L^2_{\text{ad}}([a,b] \times \Omega)$. Then the stochastic process*

$$X_t = \int_a^t f(s)\,dB(s), \quad a \le t \le b, \tag{2.10}$$

is a martingale with respect to the filtration $\{\mathcal{F}_t; a \le t \le b\}$.

We want to point out that the stochastic process X_t in Equation (2.10) being a martingale is a consequence of two facts: (1) The evaluation points of the integrand are the left endpoints of subintervals and (2) the integrand is adapted with respect to the filtration $\{\mathcal{F}_t; a \le t \le b\}$.

More generally, the Itô integral in Equation (2.7) can be extended to integrands belonging to the following space $\mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$.

Notation. Let $\mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$ denote the space of all stochastic processes $f(t,\omega)$ satisfying the conditions:

(1) $f(t)$ is *adapted* to $\{\mathcal{F}_t\}$, i.e., $f(t)$ is $\mathcal{F}_t$-measurable for each $a \le t \le b$,
(2) $\int_a^b |f(t)|^2\,dt < \infty$ almost surely.

For the precise definition of Itô integral when $f \in \mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$, see Chapter 5 of the book [14]. In this case, we do not have Theorem 2.4 because of the lack of expectation. Theorem 2.5 remains valid for $f \in \mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$. However, the corresponding theorems for Theorems 2.3 and 2.6 need modifications as follows.

Theorem 2.7. *If $f(t)$ is a continuous stochastic process and is adapted to $\{\mathcal{F}_t\}$, then $f \in \mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$ and*

$$\int_a^b f(t)\,dB(t) = \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n f(t_{i-1})\big(B(t_i) - B(t_{i-1})\big), \quad \text{in probability.}$$

Theorem 2.8. *Let $f \in \mathcal{L}_{\text{ad}}(\Omega, L^2[a,b])$. Then the stochastic process*

$$X_t = \int_a^t f(s)\,dB(s), \quad a \le t \le b,$$

is a local martingale with respect to the filtration $\{\mathcal{F}_t; a \le t \le b\}$.

For the motivation of local martingale and the proof of Theorem 2.8, see Sections 5.4 and 5.5 of the book [14].

2.5. Itô's formula. Suppose $f(x)$ is a C^2-function on the real line $\mathbb{R}$ and let $\Delta_n = \{a = t_0, t_1, \ldots, t_n = t\}$ be a partition of the internal $[a, t]$. By using the Taylor expansion we have

$$\begin{aligned} f(B(t)) - f(B(a)) &= \sum_{i=1}^{n} \Big(f(B(t_i)) - f(B(t_{i-1})) \Big) \\ &\approx \sum_{i=1}^{n} f'(B(t_{i-1}))\big(B(t_i) - B(t_{i-1})\big) \\ &\quad + \frac{1}{2} \sum_{i=1}^{n} f''(B(t_{i-1}))\big(B(t_i) - B(t_{i-1})\big)^2. \end{aligned}$$

By Theorem 2.7 we have

$$\sum_{i=1}^{n} f'(B(t_{i-1}))(B(t_i) - B(t_{i-1})) \longrightarrow \int_a^t f'(B(s))\, dB(s), \quad \text{in probability.}$$

as $\|\Delta_n\| \to 0$. On the other hand, the quadratic variation of Brownian motion in Theorem 2.1 indicates that $\big(B(t_i) - B(t_{i-1})\big)^2 \approx t_i - t_{i-1}$. Thus if the second derivative $f''(x)$ is bounded, then we can easily see that

$$\sum_{i=1}^{n} f''(B(t_{i-1}))\big(B(t_i) - B(t_{i-1})\big)^2 \longrightarrow \int_a^b f''(B(t))\, dt$$

as $\|\Delta_n\| \to 0$. Hence we have proved the following Itô's formula when $f''(x)$ is a bounded function.

Theorem 2.9. *Let $f(x)$ be a C^2-function. Then*

$$f(B(t)) = f(B(a)) + \int_a^t f'(B(s))\, dB(s) + \frac{1}{2} \int_a^t f''(B(s))\, ds. \tag{2.11}$$

The proof of Theorem 2.9, without assuming $f''(x)$ being bounded, is rather complicated. For detail, see Section 7.2 in the book [14].

It is convenient, for understanding and computation, to express the stochastic integral in Equation (2.11) in the following *stochastic differential form*:

$$df(B(t)) = f'(B(t))\, dB(t) + \frac{1}{2} f''(B(t))\, dt. \tag{2.12}$$

Notice that $dB(t)/dt$ does not exist since almost all Brownian motion paths are nowhere differentiable. Hence Equation (2.12) is understood to mean Equation (2.11) in the integral form. Moreover, recall that the chain rule for the Newton calculus is given by $\frac{d}{dt} f(g(t)) = f'(g(t))g'(t)$, or in the differential form

$$df(g(t)) = f'(g(t))\, dg(t).$$

Thus in the Itô calculus there is an extra term in Equations (2.11) and (2.12), which is due to the fact that the quadratic variation of the Brownian motion is nonzero. This extra term is often called the *Itô correction term*.

In general, suppose X_t is an *Itô process*, i.e., its stochastic differential form is given by

$$dX_t = f(t)\, dB(t) + g(t)\, dt,$$

where $f \in \mathcal{L}_{\rm ad}(\Omega, L^2[a,b])$ and g is adapted such that $\int_a^b |g(t)|\,dt < \infty$ almost surely. Let $\theta(t,x)$ be a continuous function with continuous partial derivatives θ_t, θ_x, and θ_{xx}. Then we have Itô's formula in the stochastic differential form:

$$d\theta(t,X_t) = \theta_t(t,X_t)\,dt + \theta_x(t,X_t)\,dX_t + \frac{1}{2}\theta_{xx}(t,X_t)\,(dX_t)^2, \tag{2.13}$$

where $(dX_t)^2$ is computed by using the *Itô Table* so that $(dX_t)^2 = f(t)^2\,dt$.

$\times$	$dB(t)$	dt
$dB(t)$	dt	0
dt	0	0

2.6. Stochastic differential equations. Let $\sigma(t,x)$ and $f(t,x)$ be measurable functions on $[a,b]\times\mathbb{R}$. Consider a stochastic differential equation

$$dX_t = \sigma(t,X_t)\,dB(t) + f(t,X_t)\,dt, \quad X_a = \xi, \tag{2.14}$$

where ξ is an $\mathcal{F}_a$-measurable random variable. This equation is interpreted as meaning the following stochastic integral equation

$$X_t = \xi + \int_a^t \sigma(s,X_s)\,dB(s) + \int_a^t f(s,X_s)\,dts, \quad a \le t \le b.$$

Theorem 2.10. *Suppose $\sigma(t,x)$ and $f(t,x)$ are measurable functions on $[a,b]\times\mathbb{R}$ and there exists a constant $K>0$ such that for all $t\in[a,b]$ and $x,y\in\mathbb{R}$:*

(a) $|\sigma(t,x)-\sigma(t,y)| + |f(t,x)-f(t,y)| \le K|x-y|$,
(b) $|\sigma(t,x)| + |f(t,x)| \le K(1+|x|)$.

Then the stochastic differential equation (2.14) *with $E(|\xi|^2)<\infty$ has a unique continuous solution X_t.*

Now, it is easy to derive the following conditional probability equality for a Brownian motion $B(t)$:

$$P\big(B(t)\le x \,|\, B(t_i)=y_i, i=1,2,\ldots,n\big) = P\big(B(t)\le x \,|\, B(t_n)=y_n\big). \tag{2.15}$$

for any x, y_i, $1\le i\le n$, and $0<t_1<\cdots,<t_n<t$. See Section 10.5 of the book [14] for detail.

The equality in Equation (2.15) is the motivation for the *Markov property* of a stochastic process X_t, $a\le t\le b$, i.e.,

$$P\big(X_t\le x \,|\, X_{t_i}=y_i, i=1,2,\ldots,n\big) = P\big(X_t\le x \,|\, X_{t_n}=y_n\big) \tag{2.16}$$

holds for any x, y_i, $1\le i\le n$, and $a\le t_1<t_2<\cdots<t_n<t\le b$. A stochastic process satisfying the Markov property is called a *Markov process*.

Theorem 2.11. *Under the assumption of Theorem 2.10, the solution X_t of the stochastic differential equation* (2.14) *with $E(|\xi|^2)<\infty$ is a Markov process.*

Next consider the case when the functions σ and f do not depend on t. In this case, the solution X_t, $t\ge 0$, with $X_0=x$ is a stationary Markov process (See

Theorem 10.6.2 in the book [14].) Let $\varphi(x)$ be a C^2-function. Then we can apply Itô's formula in Equation (2.13) to get

$$\begin{aligned} d\varphi(X_t) &= \varphi'(X_t)\,dX_t + \frac{1}{2}\varphi''(X_t)\,(dX_t)^2 \\ &= \varphi'(X_t)\sigma(X_t)\,dB(t) + \Big\{\frac{1}{2}\sigma(X_t)^2\varphi''(X_t) + f(X_t)\varphi'(X_t)\Big\}\,dt. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \varphi(X_t) = \varphi(x) &+ \int_0^t \varphi'(X_s)\sigma(X_s)\,dB(s) \\ &+ \int_0^t \Big\{\frac{1}{2}\sigma(X_s)^2\varphi''(X_s) + f(X_s)\varphi'(X_s)\Big\}\,ds. \end{aligned}$$

Upon taking the expectation, we get the infinitesimal generator $\mathcal{A}$ of X_t,

$$\begin{aligned} \mathcal{A}\varphi''(x) &= \lim_{t\downarrow 0}\frac{E\varphi(X_t) - \varphi(x)}{t} \\ &= E\bigg(\lim_{t\downarrow 0}\frac{1}{t}\int_0^t \Big\{\frac{1}{2}\sigma(X_s)^2\varphi''(X_s) + f(X_s)\varphi'(X_s)\Big\}\,ds\bigg) \\ &= \frac{1}{2}\sigma(x)^2\varphi''(x) + f(x)\varphi'(x). \end{aligned}$$

Hence the infinitesimal generator $\mathcal{A}$ is the differential operator

$$\mathcal{A} = \frac{1}{2}\sigma(x)^2\frac{d^2}{dx^2} + f(x)\frac{d}{dx}.$$

Conversely, suppose we are given an infinitesimal generator $\mathcal{A}$, which can be shown to be of the form

$$\mathcal{A} = \frac{1}{2}Q(x)\frac{d^2}{dx^2} + f(x)\frac{d}{dx}, \tag{2.17}$$

where $Q(x) \geq 0$ is the *diffusion coefficient* and $f(x)$ is the *drift*. Assume that there exists a constant $c > 0$ such that

$$Q(x) \geq c, \quad \forall x \in \mathbb{R},$$

and the functions Q and f satisfy the Lipschitz and linear growth conditions in Theorem 2.10. Then we can solve the following stochastic differential equation

$$dX_t = \sqrt{Q(X_t)}\,dB(t) + f(X_t)\,dt, \quad X_0 = x, \tag{2.18}$$

to get a stochastic process X_t. By the above discussion, the infinitesimal generator of X_t is the given operator $\mathcal{A}$ in Equation (2.17). This method of constructing a stochastic process directly from a given infinitesimal generator is the original motivation for K. Itô to introduce his theory of stochastic integration. For detail, see Section 10.9 of the book [14].

2.7. Multiple Wiener–Itô integrals. N. Wiener [26] introduced homogeneous chaos in 1938 in order to analyze the Hilbert space of Brownian functionals. An important property of homogeneous chaos is that homogeneous chaoses of different orders are orthogonal. In particular, all homogeneous chaoses of order $n \geq 1$ have expectation zero.

The space of homogeneous chaoses of order 1 is the Hilbert space spanned by Wiener integrals as described in Subsection 2.3. Wiener defined homogeneous chaoses of orders $n \geq 2$ by multiplying Wiener integrals and taking projections. For detail, see Section 9.4 of the book [14].

Several years after K. Itô defined the stochastic integral, he was wondering whether homogeneous chaoses can be defined in a probabilistic way by Wiener-like integrals. (*He told me in another private lecture at his home in Ithaca, New York in 1969 when I was a graduate student at Cornell University.*) Consider the case $n = 2$ with integrand $f(s)g(t)$, it seems to be natural to have the equality:

$$\int_a^b \int_a^b f(s)g(t)\, dB(s)dB(t) = \Big(\int_a^b f(s)\, dB(s)\Big)\Big(\int_a^b g(t)\, dB(t)\Big). \tag{2.19}$$

However, the right hand side of Equation (2.19) may not have expectation zero. Thus the left hand side, if given by this equation, cannot be a homogeneous chaos of order 2. Then Itô realized that the projections used by Wiener is equivalent to removing the values of an integrand on all "diagonals" meaning that at least two coordinates are equal. This led Itô to define multiple Wiener integral in 1951 [8]:

$$I_n(f) = \int_a^b \int_a^b \cdots \int_a^b f(t_1, t_2, \ldots, t_n)\, dB(t_1)dB(t_2)\cdots dB(t_n) \tag{2.20}$$

for $f \in L^2([a,b]^n)$. For detail, see Section 9.6 of the book [14]. For example, for continuous functions f and g on $[a,b]$, the definition yields the following

$$\begin{aligned}
&\int_a^b \int_a^b f(s)g(t)\, dB(s)dB(t) \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i\neq j} f(u_{i-1})g(u_{j-1})\big(B(u_i) - B(u_{i-1})\big)\big(B(u_j) - B(u_{j-1})\big) \\
&= \Big(\int_a^b f(s)\, dB(s)\Big)\Big(\int_a^b g(t)\, dB(t)\Big) - \int_a^b f(t)g(t)\, dt.
\end{aligned}$$

In particular, when $a = 0$, $b = 1$, and $f = g = 1$, we have

$$\int_0^1 \int_0^1 1\, dB(s)dB(t) = B(1)^2 - 1. \tag{2.21}$$

Now, if we perform iterated integration for this double integral, then we get

$$\int_0^1 \int_0^1 1\, dB(s)dB(t) = \int_0^1 B(1)\, dB(t). \tag{2.22}$$

However, the stochastic integral in the right hand side is not an Itô integral since the integrand $B(1)$ is not adapted to the filtration associated with the Brownian motion $B(t)$. Because of this difficulty, Itô had to cut down the region for iteration so that after each iteration the resulting integrand is adapted. But in order to compensate this subregion, the symmetrization of the integrand must be used to have the next theorem.

Theorem 2.12. *Let $f \in L^2([a,b]^n)$. Then*

$$I_n(f) = n! \int_a^b \cdots \int_a^{t_{n-2}} \left[\int_a^{t_{n-1}} \widehat{f}\,(t_1, t_2, \ldots, t_n)\, dB(t_n) \right] dB(t_{n-1}) \cdots dB(t_1),$$

where $\widehat{f}$ is the symmetrization of f.

We also mention an important property of homogeneous chaoses which shows that for each $n \geq 1$, the mapping $f \mapsto I_n(f)$ is an isometry from the space $L^2_{\text{symm}}([a,b]^n)$ into $L^2(\Omega)$ up to a scale of $\sqrt{n!}$. Here sub-"symm" means the symmetric functions.

Theorem 2.13. *Let $f \in L^2([a,b]^n)$. Then $E\big(I_n(f)\big) = 0$ and*

$$E\big(|I_n(f)|^2\big) = n! \int_a^b \int_a^b \cdots \int_a^b \big|\widehat{f}(t_1, t_2, \ldots, t_n)\big|^2\, dt_1 dt_2 \cdots dt_n.$$

3. White Noise Theory

A *white noise* can be thought of as a generalized Gaussian process $z(t)$ with mean zero and covariance given by

$$E\big(z(s)z(t)\big) = \delta_t(s),$$

where δ_t is the Dirac delta function at t. This white noise $z(t)$ is often used as a disturbance in applied mathematics.

Let $B_1(t)$ and $B_2(t)$ be independent Brownian motions with $t \geq 0$. Define

$$B(t) = \begin{cases} B_1(t), & t \geq 0, \\ B_2(-t), & t < 0. \end{cases}$$

Then $B(t)$ a Brownian motion with $t \in \mathbb{R}$. Its derivative $\dot{B}(t)$ does not exist since almost all Brownian motion paths are nowhere differentiable. But $\dot{B}(t)$ can be regarded as a generalized Gaussian process. Informally, its mean is given by

$$E(\dot{B}(t)) = \lim_{\Delta \to 0} E\Big(\frac{B(t+\Delta) - B(t)}{\Delta}\Big) = 0.$$

To derive the covariance of $\dot{B}(t)$, we need the following lemma, which can be easily verified by using integration by parts formula.

Lemma 3.1. *Let $F(t)$ be a C^1-function on $\mathbb{R}$ with compact support. Then*

$$\int_{\mathbb{R}} B(t) F'(t)\, dt = -\int_{\mathbb{R}} F(t)\, dB(t). \tag{3.1}$$

Now, let F and G be C^1-functions on $\mathbb{R}$ with compact support. Let $f = F'$ and $g = G'$. Then by using Lemma 3.1, we see that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\big[\big(B(s+\Delta_1)-B(s)\big)\big(B(t+\Delta_2)-B(t)\big)\big] f(s)g(t)\,dsdt \\
&= E\Big(\int_{-\infty}^{\infty} \big(B(s+\Delta_1)-B(s)\big) f(s)\,ds \int_{-\infty}^{\infty} \big(B(t+\Delta_2)-B(t)\big) g(t)\,dt\Big) \\
&= E\Big(\int_{-\infty}^{\infty} \big(F(s)-F(s+\Delta_1)\big)\,dB(s) \int_{-\infty}^{\infty} \big(G(t)-G(t+\Delta_2)\big)\,dB(t)\Big) \\
&= \int_{-\infty}^{\infty} \Big(F(t)\big(G(t)-G(t+\Delta_2)\big) + F(t+\Delta_1)\big(G(t+\Delta_2)-G(t)\big)\Big)\,dt.
\end{aligned}$$

Here in the last equality we have used Equation (2.3). Then we have

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\big(\dot{B}(s)\dot{B}(t)\big) f(s)g(t)\,dsdt \\
&= \lim_{\Delta_1,\Delta_2\to 0} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E\Big(\frac{B(s+\Delta_1)-B(s)}{\Delta_1}\,\frac{B(t+\Delta_2)-B(t)}{\Delta_2}\Big) f(s)g(t)\,dsdt \\
&= \lim_{\Delta_1,\Delta_2\to 0} \frac{1}{\Delta_1\Delta_2}\int_{-\infty}^{\infty} \Big(F(t)\big(G(t)-G(t+\Delta_2)\big) \\
&\qquad\qquad + F(t+\Delta_1)\big(G(t+\Delta_2)-G(t)\big)\Big)\,dt \\
&= \int_{-\infty}^{\infty} f(t)g(t)\,dt,
\end{aligned}$$

which implies that in the generalized sense

$$E\big(\dot{B}(s)\dot{B}(t)\big) = \delta_t(s).$$

Thus $\dot{B}(t)$ is a white noise. In fact, the informal derivative $\dot{B}(t)$ of a Brownian motion $B(t)$ is often taken by definition as a white noise.

Before K. Itô introduced stochastic integrals in 1942, white noise had already been used as a disturbance. Itô combined the "meaningless" white noise $\dot{B}(t)$ with infinitesimal dt to form a stochastic differential

$$\dot{B}(t)\,dt = dB(t).$$

In 1975 T. Hida [4] introduced the theory of white noise in order to define $\dot{B}(t)$ as a rigorous mathematical object. Moreover, Hida had in mind to study functionals of white noise, e.g., to treat Feynman integral as a white noise functional. Below we briefly describe the theory of white noise. For detail, see the book [13].

3.1. White noise space. Consider the real Hilbert space $L^2(\mathbb{R})$. Let

$$A = -\frac{d^2}{dx^2} + x^2 + 1. \tag{3.2}$$

The operator A is densely defined on $L^2(\mathbb{R})$ with eigenvalues $2n+2, n \geq 0$. By using this operator A, we can reconstruct the Schwartz space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing functions on $\mathbb{R}$. First use the Riesz representation theorem to identify the dual space of $L^2(\mathbb{R})$ with itself. Then we get the following triple:

$$\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R}),$$

where $\mathcal{S}'(\mathbb{R})$ is the dual space of $\mathcal{S}(\mathbb{R})$. Moreover, the Schwartz space $\mathcal{S}(\mathbb{R})$ is a nuclear space so that the above triple is a *Gel'fand triple*.

Consider the function $F : \mathcal{S}(\mathbb{R}) \longrightarrow \mathbb{C}$ given by

$$F(\xi) = \exp\Big[-\frac{1}{2}\int_{\mathbb{R}} |\xi(x)|^2\, dx\Big].$$

It is easy to check that F satisfies the following conditions:

(1) $F(0) = 1$,
(2) F is continuous,
(3) F is positive definite.

Hence we can apply the Minlos theorem to get a unique probability measure μ on the dual space $\mathcal{S}'(\mathbb{R})$ such that

$$\int_{\mathcal{S}'(\mathbb{R})} e^{i\langle x,\xi\rangle}\, d\mu(x) = e^{-\frac{1}{2}|\xi|^2}, \quad \forall\, \xi \in \mathcal{S}(\mathbb{R}),$$

where $\langle\cdot,\cdot\rangle$ is the bilinear pairing of $\mathcal{S}'(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$ and $|\xi|^2 = \int_{\mathbb{R}} |\xi(x)|^2\, dx$.

The probability space $(\mathcal{S}'(\mathbb{R}), \mu)$ is called a *white noise space*. For simplicity, we will use the notation

$$(L^2) \equiv L^2(\mathcal{S}'(\mathbb{R}), \mu). \tag{3.3}$$

For each ξ in $\mathcal{S}(\mathbb{R})$, the function $\langle\cdot,\xi\rangle$ is defined everywhere on $\mathcal{S}'(\mathbb{R})$ and is a Gaussian random variable with mean 0 and variance $|\xi|^2$. If $f \in L^2(\mathbb{R})$, we choose a sequence $\{\xi_n\}$ in $\mathcal{S}(\mathbb{R})$ such that $\xi_n \to f$ in $L^2(\mathbb{R})$. Then the sequence $\langle\cdot,\xi_n\rangle$ converges in (L^2). Define

$$\langle\cdot,f\rangle = \lim_{n\to\infty} \langle\cdot,\xi_n\rangle, \quad \text{in } (L^2).$$

It is easy to see that $\langle\cdot,f\rangle$ is well defined and is a Gaussian random variable with mean 0 and variance $|f|^2$.

It is easy to check that the following stochastic process

$$B(t,x) = \begin{cases} \langle x, 1_{[0,t)}\rangle, & t \geq 0,\, x \in \mathcal{S}'(\mathbb{R}), \\ -\langle x, 1_{[t,0)}\rangle, & t < 0,\, x \in \mathcal{S}'(\mathbb{R}), \end{cases} \tag{3.4}$$

is a Brownian motion. Informally, we have the time derivative $\dot{B}(t) = x(t)$ so that $\dot{B}$ can be regarded as an element in the white noise space $\mathcal{S}'(\mathbb{R})$.

3.2. Test and generalized functions on white noise space. Consider the space (L^2) defined in Equation (3.3). By the Wiener–Itô theorem, each $\varphi \in (L^2)$ can be uniquely represented by

$$\varphi = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L^2_{\text{symm}}(\mathbb{R}^n), \tag{3.5}$$

where I_n is the multiple Wiener–Itô integral of order n defined in Equation (2.20) with $[a,b]$ being replaced by $\mathbb{R}$ and $B(t)$ the Brownian motion given in Equation (3.4). Moreover, we have the (L^2)-norm of φ:

$$\|\varphi\| = \bigg(\sum_{n=0}^{\infty} n! |f_n|^2_{L^2(\mathbb{R}^n)}\bigg)^{1/2}. \tag{3.6}$$

For each $p \geq 0$ and φ represented by Equation (3.5), define

$$\|\varphi\|_p = \bigg(\sum_{n=0}^{\infty} n! |(A^p)^{\otimes n} f_n|^2_{L^2(\mathbb{R}^n)}\bigg)^{1/2}$$

and let

$$(\mathcal{S})_p = \big\{\varphi \in (L^2);\ \|\varphi\|_p < \infty\big\}.$$

Then we take the following *projective limit*:

$$(\mathcal{S}) = \text{the projective limit of } \big\{(\mathcal{S})_p;\ p \geq 0\big\}.$$

The space $(\mathcal{S})$ serves as a space of *test functions* defined on $\mathcal{S}'(\mathbb{R})$. Its dual space $(\mathcal{S})^*$ serves as a space of *generalized functions* defined on $\mathcal{S}'(\mathbb{R})$. It is a fact that $(\mathcal{S})$ is a nuclear space. Hence we have a Gel'fand triple

$$(\mathcal{S}) \ \subset \ (L^2) \ \subset \ (\mathcal{S})^*, \tag{3.7}$$

which is for distribution theory on the white noise space $\mathcal{S}'(\mathbb{R})$.

Here are some examples. For each ξ in $\mathcal{S}_c(\mathbb{R})$ (the complexification of $\mathcal{S}(\mathbb{R})$), the function $e^{\langle\cdot,\xi\rangle}$ is a test function. For each fixed $t \in \mathbb{R}$, the functions $\dot{B}(t)$, $:\dot{B}(t)^n:$, and $:e^{\dot{B}(t)}:$ are all generalized functions. The Feynman integral, Donsker's delta function, Kubo-Yokoi delta function are important examples of generalized functions in applications. For detail, see the book [13].

3.3. Characterization theorems. A fundamental tool in white noise theory is the S-transform. The *S-transform* of $\Phi \in (\mathcal{S})^*$ is the function

$$S\Phi(\xi) = \langle\!\langle \Phi, :e^{\langle\cdot,\xi\rangle}: \rangle\!\rangle, \quad \xi \in \mathcal{S}_c(\mathbb{R}),$$

where $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ is the bilinear pairing of $(\mathcal{S})^*$ and $(\mathcal{S})$, $\mathcal{S}_c(\mathbb{R})$ is the complexification of $\mathcal{S}(\mathbb{R})$ and $:e^{\langle\cdot,\xi\rangle}:$ is given by

$$:e^{\langle\cdot,\xi\rangle}: = \sum_{n=0}^{\infty} \frac{1}{n!} \langle :\cdot^{\otimes n}:, \xi^{\otimes n}\rangle.$$

In the early stage after T. Hida introduced the theory of white noise in 1975, it was quite hard to check whether a functional on $\mathcal{S}'(\mathbb{R})$ is actually a generalized function, in particular, one would have to find its Wiener–Itô decomposition in the generalized sense.

In 1991, Potthoff and Streit [23] did a revolutionary work for white noise theory. They proved the next theorem to characterize generalized functions in the space $(\mathcal{S})^*$ in terms of the analytic and growth conditions of its S-transform.

Theorem 3.2. *A function $F : \mathcal{S}_c(\mathbb{R}) \longrightarrow \mathbb{C}$ is the S-transform of $\Phi \in (\mathcal{S})^*$ if and only if it satisfies the following conditions:*

(1) *For any $\xi, \eta \in \mathcal{S}_c(\mathbb{R})$, the function $F(z\xi+\eta)$ is an entire function of $z \in \mathbb{C}$;*

(2) *There exist constants* $K, a, p > 0$ *such that*

$$|F(\xi)| \leq K \exp\Big[a \int_{\mathbb{R}} \big|(A^p \xi)(x)\big|^2 \, dx\Big], \quad \forall\, \xi \in \mathcal{S}_c(\mathbb{R}),$$

where A *is the operator in Equation* (3.2).

For the white noise approach to stochastic integration, we use the basic Gel'fand triple in Equation (3.7) and the above characterization theorem. We mention that there are many other Gel'fand triples and the corresponding characterization theorems in the work of L. Streit and his collaborators, Y.-J. Lee, Ouerdiane and his collaborators, my joint work with W. G. Cochran and A. N. Sengupta, and with N. Asai and I. Kubo.

3.4. White noise differentiation operator. Let $h \in \mathcal{S}'(\mathbb{R})$ and $\varphi \in (\mathcal{S})$. The *Gâteaux derivative* of φ in the direction h is given by

$$D_h \, \varphi(x) = \lim_{\epsilon \to 0} \frac{\varphi(x + \epsilon h) - \varphi(x)}{\epsilon}.$$

The operator D_h can be shown to be a continuous operator from $(\mathcal{S})$ into itself. Hence its adjoint D_h^* is a continuous operator from $(\mathcal{S})^*$ into itself.

In particular, let $h = \delta_t$, the Dirac delta function at $t \in \mathbb{R}$. Then we have the *white noise differentiation operator* ∂_t:

$$\partial_t \equiv D_{\delta_t}.$$

Thus we have continuous operators $\partial_t : (\mathcal{S}) \longrightarrow (\mathcal{S})$ and $\partial_t^* : (\mathcal{S})^* \longrightarrow (\mathcal{S})^*$ for each $t \in \mathbb{R}$.

Let ν be the standard Gaussian measure on $\mathbb{R}^n$. By the integration by parts formula, we have the equality

$$\int_{\mathbb{R}^n} \big(D_h f(x)\big) g(x) \, d\nu(x) = \int_{\mathbb{R}^n} f(x) \big\{ g(x) \langle x, h\rangle - D_h g(x) \big\} \, d\nu(x)$$

for suitable functions f and g. The infinite dimensional analogue of this formula for the white noise space $(\mathcal{S}'(\mathbb{R}), \mu)$ is given by

$$\int_{\mathcal{S}'(\mathbb{R})} \big(D_h \varphi(x)\big) \psi(x) \, d\mu(x) = \int_{\mathcal{S}'(\mathbb{R})} \varphi(x) \big\{ \psi(x) \langle x, h\rangle - D_h \psi(x) \big\} \, d\mu(x),$$

where $h \in \mathcal{S}'(\mathbb{R})$ and $\varphi, \psi \in (\mathcal{S})$. This equality yields the adjoint operator of D_h, i.e., $D_h^* = \langle \cdot, h\rangle - D_h$, which can be rewritten as

$$\langle \cdot, h\rangle = D_h + D_h^*$$

as continuous operators from the space $(\mathcal{S})$ into $(\mathcal{S})^*$. Take $h = \delta_t$, the Dirac delta function at $t \in \mathbb{R}$ and note that $\langle \cdot, \delta_t\rangle = \dot{B}(t)$. Then we have the equality

$$\dot{B}(t) = \partial_t + \partial_t^*. \tag{3.8}$$

Therefore, for each $t \in \mathbb{R}$, we can view the quantity $\dot{B}(t)$ in two ways:

(a) $\dot{B}(t)$ is a generalized function in $(\mathcal{S})^*$ and $\big(S\dot{B}(t)\big)(\xi) = \xi(t)$.

(b) $\dot{B}(t)$, as a multiplication operator, is continuous from $(\mathcal{S})$ into $(\mathcal{S})^*$ and Equation (3.8) holds.

3.5. White noise integral. Consider the Brownian motion $B(t)$, $t \geq 0$, defined in Equation (3.4). Let $\{\mathcal{F}_t\}$ be the filtration with $\mathcal{F}_t = \sigma\{B(s), 0 \leq s \leq t\}$. In the Itô calculus, the white noise $\dot{B}(t)$ and time differential dt are combined together to form the stochastic differential

$$\dot{B}(t)\, dt = dB(t).$$

But now, with white noise theory, we know that the white noise $\dot{B}(t)$ for each t can be regarded as a generalized function in $(\mathcal{S})^*$ or as a continuous operator $\dot{B}(t) = \partial_t + \partial_t^*$ from $(\mathcal{S})$ into $(\mathcal{S})^*$. It is natural to ask whether we can interpret an Itô integral as

$$\int_a^b f(t)\, dB(t) = \int_a^b \dot{B}(t) f(t)\, dt \ ? \tag{3.9}$$

$$\int_a^b f(t)\, dB(t) = \int_a^b (\partial_t + \partial_t^*) f(t)\, dt \ ? \tag{3.10}$$

Note that the integrand in the right hand side of Equation (3.9) makes sense as a generalized function in the space $(\mathcal{S})^*$ only when $f(t) \in (\mathcal{S})$ for each t. Hence this equation does not fit to the Itô calculus. On the other hand, the white noise differentiation ∂_t in Equation (3.10), when applied to a stochastic process $f(t)$, is not well defined. In fact, we need to consider the right-hand and left-hand white noise differentiation operators ∂_{t+} and ∂_{t-}. For detail, see the book [13].

This leads to the integral in Equation (3.10) with the operator ∂_t^*. Actually, I. Kubo and S. Takenaka [11] obtained the next theorem in 1981. Recall that we are using the Brownian motion $B(t)$ in Equation (3.4) and its associated filtration $\mathcal{F}_t = \sigma\{B(s), 0 \leq s \leq t\}$.

Theorem 3.3. *Let $f(t,x)$, $a \leq t \leq b, x \in \mathcal{S}'(\mathbb{R})$, be an adapted stochastic process such that $\int_a^b \int_{\mathcal{S}'(\mathbb{R})} |f(t,x)|^2\, d\mu(x) dt < \infty$. Then*

$$\int_a^b f(t)\, dB(t) = \int_a^b \partial_t^* f(t)\, dt, \tag{3.11}$$

where the left hand side is an Itô integral.

For the proof of this theorem, see Theorem 13.12 in the book [13]. Thus an Itô integral of a stochastic process $f \in L^2_{\text{ad}}([a,b] \times \mathcal{S}'(\mathbb{R}))$ can be expressed as a white noise integral. However, for $f \in \mathcal{L}_{\text{ad}}(\mathcal{S}'(\mathbb{R}), L^2[a,b])$, we cannot express its Itô integral as a white noise integral.

3.6. Hitsuda–Skorokhod integral. Observe that the white noise integral in the right hand side of Equation (3.11) does not require the stochastic process $f(t)$ to be adapted with respect to the Brownian motion filtration. Hence the white noise integral $\int_a^b \partial_t^* f(t)\, dt$ provides an extension of the Itô integral to stochastic processes $f(t)$ which may not be adapted, i.e., possibly anticipating. For example, we have the white noise integral

$$\int_0^1 \partial_t^* B(1)\, dt = B(1)^2 - 1.$$

(See Example 13.14 in the book [13].) Hence we have an extension of the following Itô integral

$$\int_0^1 B(1)\, dB(t) = B(1)^2 - 1, \tag{3.12}$$

where the integrand $B(1)$ is not adapted with respect to the Brownian motion $B(t)$ for $0 \leq t \leq 1$.

There is a theorem due to N. Obata [22] stating that

$$\bigcup_{p>1} L^p(\mathcal{S}'(\mathbb{R}), \mu) \;\subset\; (\mathcal{S})^*.$$

(See also Section 8.5 in the book [13].) Being motivated by this fact, we say that the white noise integral

$$\int_a^b \partial_t^* f(t)\, dt \tag{3.13}$$

is the *Hitsuda–Skorokhod integral* of f if it is represented by a random variable in $L^p(\mathcal{S}'(\mathbb{R}), \mu)$ for some $p > 1$, i.e., there exists $\Phi \in L^p(\mathcal{S}'(\mathbb{R}), \mu)$ for some $p > 1$ such that

$$\langle\!\langle \Phi,\, \varphi \rangle\!\rangle = \left\langle\!\!\left\langle \int_a^b \partial_t^* f(t)\, dt,\; \varphi \right\rangle\!\!\right\rangle, \quad \forall\, \varphi \in (\mathcal{S}).$$

The Hitsuda–Skorokhod integral in Equation (3.13), with different notation, was introduced independently by M. Hitsuda [5] in 1972 using complex Brownian motion and by A. V. Skorokhod in 1975 using Wiener–Itô theorem. (*I learned from Yu. L. Daletskii in 1989 during the Fifth International Vilnius Conference on Probability Theory and Mathematical Statistics that it was Hitsuda who first introduced this stochastic integral in 1972 at the Second Japan-USSR Symposium on Probability Theory. Daletskii also mentioned that Skorokhod attended Hitsuda's lecture and must be aware of this then new stochastic integral.*)

Example 3.4. Consider the white noise formulation of the stochastic differential equation in Equation (1.4) with a different anticipating initial condition

$$dX_t = \partial_t^* X_t\, dt, \quad X_0 = \operatorname{sgn}\big(B(1)\big), \quad 0 \leq t \leq 1. \tag{3.14}$$

It is shown in Example 13.30 in the book [13] that the solution is given by

$$X_t = \operatorname{sgn}\big(B(1) - t\big) e^{B(t) - \frac{1}{2}t}, \quad 0 \leq t \leq 1.$$

It is interesting to compare this solution X_t with the solution Y_t of Equation (1.4) given by Equation (1.8).

The next theorem is an anticipative Itô's formula obtained by Hitsuda [5] in 1972. For the proof, see Theorem 13.19 in the book [13].

Theorem 3.5. *Let $\theta(x, y)$ be a C^2-function on $\mathbb{R}^2$ such that*

$$\theta\big(B(\cdot), B(b)\big), \;\; \theta_{xx}\big(B(\cdot), B(b)\big), \;\; \theta_{xy}\big(B(\cdot), B(b)\big)$$

are all in $L^2\big([a,b];(L^2)\big)$*. Then the following equality holds in* (L^2)*:*

$$\theta\big(B(t),B(b)\big)=\theta\big(B(a),B(b)\big)+\int_a^t \partial_s^*\Big(\theta_x\big(B(s),B(b)\big)\Big)\,ds$$

$$+\int_a^t \Big(\frac{1}{2}\theta_{xx}\big(B(s),B(b)\big)+\theta_{xy}\big(B(s),B(b)\big)\Big)\,ds,\quad a\le t\le b, \tag{3.15}$$

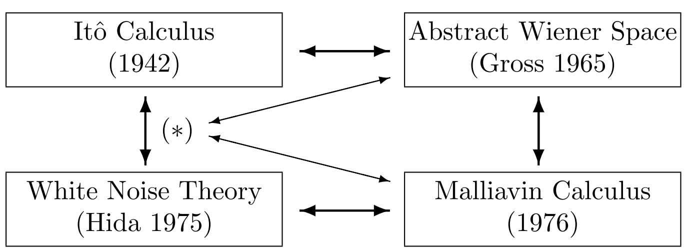
where the first integral is a Hitsuda–Skorokhod integral.

We point out that in Equation (3.15) there is an additional correction term involving θ_{xy} due to the anticipating $B(b)$ in the y variable.

4. General Stochastic Integration

H. Lebesgue introduced measure theory around 1902-04. However, the first really significant application of the Lebesgue measure theory was given years later by N. Wiener in his construction of Brownian motion in 1923 [25]. The Wiener integral was generalized to the Itô integral by K. Itô in 1942 [6] as we explained in Section 2. On the other hand, Wiener's work inspired L. Gross to introduce abstract Wiener space in 1965 [3] for infinite dimensional analysis (see also the book [12].) Then came the white noise theory initiated by T. Hida in 1975 [4] as we described in Section 3. In his lecture for the 1976 Kyoto International Symposium, P. Malliavin [20] achieved a triumph to provide a probabilistic proof of the existence of transition probabilities (see also the book [21].)

The following diagram shows the relationships among the above four areas of stochastic analysis.



We are interested in extending Itô's theory of stochastic integration to stochastic integrals with integrands being possibly not adapted with respect to the underlined filtration. For example, recall the stochastic differential equation in Equation (1.4). Since the initial condition $Y_0 = B(1)$ is anticipating, this equation is not within the scope of Itô's theory. More generally, suppose we are given an infinitesimal generator in Equation (2.17). Then we can consider the stochastic differential equation in Equation (2.18) with an anticipating initial condition. Then we need to extend Itô's theory in order to handle this equation.

The Hitsuda–Skorokhod integral provides an extension (actually only partial extension) of Itô integral. However, within white noise theory, there are several difficulties to handle this extension:

(a) In general, the integral $\int_a^b \partial_t^* f(t)\,dt$ of a stochastic process $f(t)$ is a white noise integral that defines a generalized function in the space $(\mathcal{S})^*$ and has no probabilistic meaning.
(b) The white noise integral $\int_a^b \partial_t^* f(t)\,dt$ is a Hitsuda–Skorokhod integral if it is represented by a random variable in the space $L^p(\mathcal{S}'(\mathbb{R}), \mu)$ for some $p > 1$. However, there is no theorem which asserts when a generalized function in $(\mathcal{S})^*$ belongs to $L^p(\mathcal{S}'(\mathbb{R}), \mu)$, $p > 1$.

We want to follow Itô's ideas in his lecture for the 1976 International Symposium [9], yet keep in mind white noise methods, to give a real extension of the Itô integral. We expect this extension to be connected to abstract Wiener space theory and the Malliavin calculus.

4.1. Motivating ideas. Let $B(t)$ be a Brownian motion and take $\{\mathcal{F}_t\}$ to be the filtration, i.e., $\mathcal{F}_t = \sigma\{B(s);\, 0 \le s \le t\}$. At the 1976 Kyoto International Symposium, K. Itô started his lecture [9] by writing on the blackboard the following question (*I was in the audience*):

$$\int_0^1 B(1)\,dB(t) = B(1)\int_0^1 dB(t) = B(1)^2 \ \ ? \tag{4.1}$$

Then he pointed out that $\int_0^1 B(1)\,dB(t)$ cannot be defined as a stochastic integral with respect to a Brownian motion since the integrand $B(1)$ is not adapted with respect to $B(t)$. To overcome this difficulty, he enlarged the filtration

$$\mathcal{G}_t = \alpha\{B(1), B(s);\, 0 \le s \le t\}.$$

and decomposed the Brownian motion $B(t)$ as

$$B(t) = \left(B(t) - \int_0^t \frac{B(1)-B(u)}{1-u}\,du\right) + \int_0^t \frac{B(1)-B(u)}{1-u}\,du.$$

so that $B(t)$ is a quasimartigale. Obviously, the integrand $B(1)$ is adapted with respect to $\{\mathcal{G}_t\}$. Thus $\int_0^1 B(1)\,dB(t)$ is defined as a stochastic integral with respect to a quasimartingale and Equation (4.1) holds. In general,

$$(\text{Itô})\int_0^t B(1)\,dB(s) = B(1)B(t), \quad 0 \le t \le 1. \tag{4.2}$$

Observe that when $t = 1$, this integral is different from the Hitsuda–Skorokhod integral in Equation (3.12) as defined through white noise theory.

We notice that the essence of Itô's ideas is the following two points:

(a) Keep the integrand $B(1)$.
(b) Change the filtration and decompose the integrator $B(t)$.

Being inspired by Itô's ideas, we would try to reverse the roles of the integrand and the integrator, namely,

(1) Keep the Brownian motion $B(t)$ and the filtration.
(2) Decompose the integrand $B(1)$.

The integrand $B(1)$ is decomposed as follows:

$$B(1) = \Big(B(1) - B(t)\Big) + B(t). \tag{4.3}$$

Let $\Delta_n = \{0 = s_0, s_1, \ldots, s_{n-1}, s_n = t\}$ be a partition of the interval $[0, t]$ with fixed t in $[0, 1]$. On the subinterval $[s_{i-1}, s_i]$, we evaluate the first term $B(1) - B(t)$ in Equation (4.3) at the right endpoint s_i. On the other hand, we evaluate the second term $B(t)$ (which is adapted) at the left endpoint s_{i-1} just as in an Itô integral. Then we have

$$\begin{aligned}
&\int_0^t B(1)\, dB(s) \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n \Big\{ \Big(B(1) - B(s_i) \Big) + B(s_{i-1}) \Big\} \big(B(s_i) - B(s_{i-1}) \big) \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n \Big\{ B(1) - \Big(B(s_i) - B(s_{i-1}) \Big) \Big\} \big(B(s_i) - B(s_{i-1}) \big) \\
&= B(1)B(t) - \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n \big(B(s_i) - B(s_{i-1}) \big)^2 \\
&= B(1)B(t) - t,
\end{aligned}$$

where in the last equality we have used the quadratic variation of the Brownian motion $B(t)$. Therefore, we have

$$\int_0^t B(1)\, dB(s) = B(1)B(t) - t, \quad 0 \le t \le 1. \tag{4.4}$$

Observe that this new stochastic integral in Equation (4.4) is different from the one in Equation (4.2), but agrees with the Hitsuda–Skorokhod integral in Equation (3.12) when $t = 1$.

To illustrate our ideas, we do one more example to evaluate the integral

$$\int_0^t B(1)B(s)\, dB(s), \quad 0 \le t \le 1.$$

The integrand is decomposed in terms of $B(s)$ and $B(1) - B(s)$ as follows:

$$B(1)B(s) = \big(B(1) - B(s) \big) B(s) + B(s)^2.$$

On the subinterval $[s_{i-1}, s_i]$, we evaluate $B(s)$ and $B(1) - B(s)$ at the left endpoint and right endpoint, respectively. Then we get

$$\begin{aligned}
&\int_0^t B(1)B(s)\, dB(s) \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n \Big\{ \Big(B(1) - B(s_i) \Big) B(s_{i-1}) + B(s_{i-1})^2 \Big\} \big(B(s_i) - B(s_{i-1}) \big) \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n \Big\{ B(1)B(s_{i-1}) - B(s_{i-1}) \Big(B(s_i) - B(s_{i-1}) \Big) \Big\} \big(B(s_i) - B(s_{i-1}) \big) \\
&= B(1) \int_0^t B(s)\, dB(s) - \int_0^t B(s)\, ds \\
&= \frac{1}{2} B(1) \big(B(t)^2 - t \big) - \int_0^t B(s)\, ds,
\end{aligned}$$

where in the third equality we used the fact that $\big(B(s_i) - B(s_{i-1})\big)^2 \approx s_i - s_{i-1}$. Hence we have the equality

$$\int_0^t B(1)B(s)\,dB(s) = \frac{1}{2}B(1)\big(B(t)^2 - t\big) - \int_0^t B(s)\,ds, \quad 0 \le t \le 1. \tag{4.5}$$

4.2. Definition of a new stochastic integral. In view of the discussion in the previous subsection 4.1, we see that the Itô theory has a counterpart consisting of stochastic processes that are not adapted, but have a very special property. When those stochastic processes are taken as integrands, the evaluation points are the right endpoints of subintervals.

As in the Itô theory, we fix a Brownian motion $B(t)$ and a filtration $\{\mathcal{F}_t\}$ such that (1) $B(t)$ is $\mathcal{F}_t$-measurable for each t and (2) $B(t) - B(s)$ and $\mathcal{F}_s$ are independent for any $s \le t$.

Definition 4.1. A stochastic process $\varphi(t)$ is called *instantly independent* with respect to a filtration $\{\mathcal{F}_t\}$ if $\varphi(t)$ and $\mathcal{F}_t$ are independent for each t.

We will refer the set of adapted stochastic processes as the Itô part and the set of instantly independent stochastic processes as the counterpart. It is easy to see that if a stochastic process $g(t,\omega)$ is both adapted and instantly independent, then it must be a deterministic function $g(t)$ of t.

Definition 4.2. Let $f(t)$ be a continuous adapted stochastic process and $\varphi(t)$ a continuous instantly independent stochastic process. Define the *new stochastic integral* of $f(t)\varphi(t)$ by

$$\int_a^b f(t)\varphi(t)\,dB(t) = \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n f(t_{i-1})\varphi(t_i)\big(B(t_i) - B(t_{i-1})\big), \tag{4.6}$$

provided that the limit in probability exists.

We use the linearity to extend this stochastic integral to stochastic processes which are finite linear combinations of such products $f(t)\varphi(t)$.

We quickly point out the relationship of this new stochastic integral and multiple Wiener–Itô integral. Consider the simple example

$$\int_0^1 \int_0^1 1\,dB(s)dB(t).$$

Suppose we apply iterated integration. Then we can use Equation (4.4) to get

$$\int_0^1 B(1)\,dB(t) = B(1)^2 - 1,$$

which is exactly the double Wiener–Itô integral of 1. Hence we have

$$\int_0^1 \int_0^1 1\,dB(s)dB(t) = \int_0^1 \Big(\int_0^1 1\,dB(s)\Big)dB(t).$$

More generally, we can use the new stochastic integral to evaluate multiple Wiener–Itô integrals by performing iterated integration (which makes sense with our new stochastic integration), exactly like in the ordinary calculus. This property seems to be interesting for stochastic analysis. For detail, see [2].

4.3. Near martingale. In the Itô theory of stochastic integration, two guiding motivations are the martingale property and Markov property, which lead to the evaluation points of an integrand at the left endpoints of subintervals.

With our evaluation points in Definition 4.2, we certainly do not have martingale property. But then what kind of property do we have? Consider the stochastic process in Equation (4.4), i.e.,

$$X_t = \int_0^t B(1)\,dB(s) = B(1)B(t) - t, \quad 0 \le t \le 1. \tag{4.7}$$

Recall that we have an underlined filtration $\{\mathcal{F}_t\}$. Let $s \le t$. Then we have

$$\begin{aligned} E(X_t|\,\mathcal{F}_s) &= E(B(1)B(t) - t|\,\mathcal{F}_s) \\ &= E\Big(\big\{[B(1) - B(t)] + [B(t) - B(s)] + B(s)\big\} \\ &\qquad \times \big\{[B(t) - B(s)] + B(s)\big\}\Big|\,\mathcal{F}_s\Big) - t \\ &= t - s + B(s)^2 - t \\ &= B(s)^2 - s, \end{aligned} \tag{4.8}$$

which is clearly not equal to X_s almost surely. Hence X_t is not a martingale with respect to the filtration $\{\mathcal{F}_t\}$. However, observe that X_s is not $\mathcal{F}_s$-measurable. Moreover, by putting $t = s$ in Equation (4.8), we get

$$E(X_s|\,\mathcal{F}_s) = B(s)^2 - s. \tag{4.9}$$

It follows from Equations (4.8) and (4.9) that for any $s \le t$,

$$E(X_t|\,\mathcal{F}_s) = E(X_s|\,\mathcal{F}_s), \quad \text{a.s.}$$

This equality is the motivation for the concept in the next definition.

Definition 4.3. A stochastic process X_t is called a *near-martingale* with respect to a filtration $\{\mathcal{F}_t\}$ if for any $s \le t$ it holds that

$$E(X_t|\,\mathcal{F}_s) = E(X_s|\,\mathcal{F}_s), \quad \text{a.s.} \tag{4.10}$$

or equivalently,

$$E(X_t - X_s|\,\mathcal{F}_s) = 0, \quad \text{a.s.}$$

Note that if a near-martingale X_t is adapted, then it is a martingale. Moreover, by taking expectation in both sides of Equation (4.10), we see that $EX_t = EX_s$ for any $s \le t$. Hence near-martingale implies fair game.

The next theorem shows that the near-martingale property is the analogue of martingale property in the Itô integral for the new stochastic integral that we have defined. For detail, see [17].

Theorem 4.4. *Let f and φ be continuous functions on $\mathbb{R}$ and*

$$X_t = \int_a^t f\big(B(s)\big)\varphi\big(B(b) - B(s)\big)\,dB(s), \quad a \le t \le b.$$

Assume that $E|X_t| < \infty$ for all $t \in [a,b]$. Then the stochastic process X_t, $t \in [a,b]$, is a near-martingale.

4.4. Itô's formula. Let $\theta(x)$ be a C^2-function of x. Then by Itô's formula,

$$d\theta(B(t)) = \theta'(B(t))\, dB(t) + \frac{1}{2}\theta''(B(t))\, dt, \quad a \le t \le T,$$

which means the following equality in a stochastic integral form:

$$\theta(B(t)) = \theta(B(a)) + \int_a^t \theta'(B(s))\, dB(s) + \frac{1}{2}\int_0^t \theta''(B(s))\, ds, \quad a \le t \le T.$$

Now, suppose we have a function $\theta(x, y)$ and consider $\theta\big(B(t), B(T)\big)$. What is it equal to? In order to find out the answer, let us assume that $\theta(x, y) = f(x)\varphi(y-x)$ with f and g being C^1-functions. Then we have

$$\begin{aligned}
&\sum_{i=1}^n \theta\big(B(s_{i-1}), B(T)\big)\big(B(s_i) - B(s_{i-1})\big) \\
&= \sum_{i=1}^n f\big(B(s_{i-1})\big)\, \varphi\big(B(T) - B(s_{i-1})\big)\big(B(s_i) - B(s_{i-1})\big) \\
&\approx \sum_{i=1}^n f\big(B(s_{i-1})\big)\Big\{\varphi\big(B(T) - B(s_i)\big) + \varphi'\big(B(T) - B(s_i)\big)\big(B(s_i) - B(s_{i-1})\big)\Big\} \\
&\qquad\qquad \times \big(B(s_i) - B(s_{i-1})\big) \\
&\to \int_a^t f\big(B(s)\big)\, \varphi\big(B(T) - B(s)\big)\, dB(s) + \int_a^t f\big(B(s)\big)\, \varphi'\big(B(T) - B(s)\big)\, ds.
\end{aligned}$$

Note that $f(x)\varphi'(y-x) = \frac{\partial\theta}{\partial y}(x, y)$. Hence we have the next lemma from [1].

Lemma 4.5. *Let $f(x)$ be a continuous function and $\varphi(x)$ a C^1-function. Let $\theta(x, y) = f(x)\varphi(y - x)$. Then*

$$\begin{aligned}
&\sum_{i=1}^n \theta\big(B(s_{i-1}), B(T)\big)\big(B(s_i) - B(s_{i-1})\big) \\
&\longrightarrow \int_a^t \theta\big(B(s), B(T)\big)\, dB(s) + \int_a^t \frac{\partial\theta}{\partial y}\big(B(s), B(T)\big)\, ds, \qquad (4.11)
\end{aligned}$$

in probability as $\|\Delta_n\| \to 0$. Here Δ_n's are partitions of $[a, t]$.

Notice the second integral in Equation (4.11). This is a correction term due to the anticipating $B(T)$ in the function θ. By using this lemma, we can derive the following Itô's formula from [1].

Theorem 4.6. *Let f and φ be C^2-functions and let $\theta(x, y) = f(x)\varphi(y-x)$. Then*

$$\begin{aligned}
\theta\big(B(t), B(T)\big) &= \theta\big(B(a), B(T)\big) + \int_a^t \frac{\partial\theta}{\partial x}\big(B(s), B(T)\big)\, dB(s) \\
&+ \int_a^t \Big\{\frac{1}{2}\frac{\partial^2\theta}{\partial x^2}\big(B(s), B(T)\big) + \frac{\partial^2\theta}{\partial x \partial y}\big(B(s), B(T)\big)\Big\}\, ds, \quad a \le t \le T.
\end{aligned}$$

The above theorem provides an anticipative Itô's formula for a special case. For further generalizations, see the references [15] and [18].

4.5. Girsanov theorem. Suppose $B(t)$ is a Brownian motion with respect to a probability measure P. A very important stochastic process in the Itô calculus is the *exponential process* defined by

$$\mathcal{E}_h(t) = \exp\Big[\int_0^t h(s)\,dB(s) - \frac{1}{2}\int_0^t h(s)^2\,ds\Big], \quad 0 \le t \le T, \tag{4.12}$$

where $h(t)$ is an adapted stochastic process such that $\int_0^T |f(t)|^2\,dt < \infty$ almost surely. Consider the translation of $B(t)$ by the integral of $h(t)$, i.e.,

$$W(t) = B(t) - \int_0^t h(s)\,ds, \quad 0 \le t \le T. \tag{4.13}$$

The well-known Girsanov theorem says that if $E_P(\mathcal{E}_h(t)) = 1$ for all $t \in [0,T]$, then the stochastic process $W(t)$ in Equation (4.13) is a Brownian motion with respect to the probability measure $dQ = \mathcal{E}_h(T)\,dP$.

If one examines the proof of the Girsanov theorem (e.g., see the book [14]), one sees that the conclusion "$W(t)$ is a *Brownian motion*" can be replaced by "$W(t)$ is a *martingale*" since the other properties (those needed for applying the Lev́y characterization theorem of Brownian motion) are automatically satisfied. With this modification of the Girsanov theorem, we can extend the theorem to the new stochastic integral.

The next theorem [15] gives a special case of anticipative Girsanov theorem. See [16] for further generalizations.

Theorem 4.7. *Let $\varphi \in L^2(\mathbb{R})$. Then the stochastic process*

$$W(t) = B(t) - \int_0^t \varphi\big(B(T) - B(s)\big)\,ds, \quad 0 \le t \le T,$$

is a near-martingale with respect to the probability measure given by

$$dQ = \exp\Big\{\int_0^T \phi\big(B(T) - B(s)\big)\,dB(s) - \frac{1}{2}\int_0^T \phi^2\big(B(T) - B(s)\big)\,ds\Big\}\,dP.$$

4.6. Some remarks and open problems. By taking $\varphi \equiv 1$ in Equation (4.6), we see that our stochastic integral reduces to the Itô integral. Hence our stochastic integral is obviously a true extension of the Itô integral. However, our extension is still in the early stage and has many difficulties to be overcome. Below we mention some remarks and open problems.

1. The Itô integral $\int_a^b f(t)\,dB(t)$ is defined for stochastic processes f in the space $\mathcal{L}_{\mathrm{ad}}(\Omega, L^2[a,b])$, i.e., $f(t)$ is adapted and $\int_a^b |f(t)|^2\,dt < \infty$ almost surely. Let $f \in \mathcal{L}_{\mathrm{ad}}(\Omega, L^2[a,b])$. Then we can ask the question:

   "What is the class of stochastic processes $\varphi(t)$ for which the new stochastic integral $\int_a^b f(t)\varphi(t)\,dB(t)$ exists?"

   More generally, we need to find the class of all stochastic processes $\Phi(t)$ for which the new stochastic integral $\int_a^b \Phi(t)\,dB(t)$ exists.

2. The Hitsuda–Skorokhod integral $\int_a^b \partial_t^* f(t)\, dB(t)$ is a partial extension of the Itô integral for non-adapted stochastic processes $f(t)$. It is natural to ask the question:

   "Is the new stochastic integral related to the Hitsuda–Skorokhod integral?"

   We conjecture that when both integrals exist, they have the same value.
3. In Definition 4.3 we have the concept of near-martingale. Obviously, we can define *near-submartingale* and *near-supermartingale* by replacing the equality sign = in Equation (4.10) with $\geq$ and $\leq$, respectively. In order to prove the continuity property of a stochastic process $X_t = \int_a^t \Phi(s)\, dB(s)$ associated with a nonadapted integrand $\Phi(t)$ of our kind, we will need to obtain the *Doob inequality for near-submartingales*. We will also need the *Doob–Meyer decomposition theorem for near-submartingales*.
4. In Theorem 4.6 we stated a very simple form of Itô's formula for the new integral. This theorem has been generalized further in the papers [15] and [18]. However, the more general form of Itô's formula is yet to be derived. Moreover, in view of the relationship between Itô's formula and the Doob–Meyer decomposition in the Itô calculus, we can ask the question:

   "Is there a relationship between the yet to be derived Itô's formula and the yet to be discovered Doob–Meyer decomposition for the new integral?"

5. In the Itô calculus the exponential process of $h(t)$ is defined by Equation (4.12), which is in fact the renormalization of $\exp\big[\int_0^t h(s)\, dB(s)\big]$. When $h(t)$ is in the counterpart given by $h(t) = \phi\big(B(T) - B(t)\big)$, the exponential process takes the same form as in the Itô calculus and is given in Theorem 4.7. In general, we can ask the following question:

   "What is the exponential process associated with a non-adapted stochastic process $h(t)$?"

   Here is an example. Let θ be a C^2-function on $\mathbb{R}$. By direct computation, we can derive the following exponential process associated with $\theta(B(1))$

$$X_t = \exp\bigg\{\int_0^t \theta(B(1))\, dB(s) - \int_0^t \Big[\frac{1}{2}\theta(B(1))^2 + \Big(\theta'(B(1))B(s) - \theta''(B(1))s\Big)\theta(B(1))\Big]\, ds\bigg\}$$

   in the sense that X_t, $0 \leq t \leq 1$, is a near-martingale.
6. Special cases of the Girsanov theorem for the new stochastic integral have been obtained in the papers [15] and [16]. It would be interesting to find the *formulation of the Girsanov theorem for the general case* where the translation involves both the Itô part (adapted) and the counterpart (instantly independent). This formulation needs to use the exponential process mentioned in the previous item of remark.
7. In Section 1 we have a simple stochastic differential equation (1.4) with an anticipating initial condition. The solution is given by Equation (1.8) for $0 \leq t \leq 1$. In fact, we can also use Itô's formula in Theorem 4.6 to derive

the solution. In the paper [10] we have solved a class of linear stochastic differential equations with anticipating initial conditions. However, we have not yet been able to obtain a theorem on the existence and uniqueness of a solution of a general stochastic differential equation with the new stochastic integral. On the other hand, below is an interesting problem:

"Let X_t and Y_t be the solutions of the stochastic differential equations:

$$dX_t = f(X_t)\, dB(t) + g(X_t)\, dt, \quad X_0 = x, \quad 0 \le t \le T,$$
$$dY_t = f(Y_t)\, dB(t) + g(Y_t)\, dt, \quad X_0 = \xi(B(T)), \quad 0 \le t \le T.$$

Find the relationship between the solutions X_t and Y_t."

8. We know that the martingale property in the Itô theory of stochastic integration has an analogue, namely, the near-martingale property for the new stochastic integration. Here is the question for another important property:

   "What is the analogue of the Markov property for the counterpart of the new stochastic integration?"

9. In the paper [16] we give an application of the new integral to a simple model of the Black–Scholes equation. It would be interesting to use this model together with the classical Black–Scholes model to investigate the influence of the insider information on a market. Moreover, we hope that the new stochastic integration will be useful to study finance whenever there is insider information, which is anticipating.
10. Let (i, H, B) be an abstract Wiener space and μ the standard Gaussian measure on B [12]. Suppose $T : B \to B$ is a nonlinear transformation such that $\mu \circ T^{-1}$ is absolutely continuous with respect to μ. It is a well-known fact that the Radon-Nikodym derivative $d(\mu \circ T^{-1})/d\mu$ is related to the exponential process in the Girsanov theorem. This leads to the question:

    "Can we use the transformation formula for an abstract Wiener space to derive the exponential process needed for the Girsanov theorem for the new stochastic integral?"

    We do not have an answer yet to this question.

References

1. Ayed, W. and Kuo, H.-H.: An extension of the Itô integral, *Communications on Stochastic Analysis* **2**, no. 3 (2008) 323–333.
2. Ayed, W. and Kuo, H.-H.: An extension of the Itô integral: toward a general theory of stochastic integration, *Theory of Stochastic Processes* **16(32)**, no. 1 (2010) 1–11.
3. Gross, L.: Abstract Wiener spaces, in: *Proc. 5th Berkeley Symp. Math. Stat. and Probab.* **2**, part 1 (1965) 31–42, University of California Press, Berkeley.
4. Hida, T.: *Analysis of Brownian Functionals*, Carleton Mathematical Lecture Notes **13** 1975.
5. Hitsuda, M.: Formula for Brownian partial derivatives, *Second Japan-USSR Symp. Probab. Th.***2** (1972) 111–114.
6. Itô, K.: Differential equations determining Markov processes, *Zenkoku Shijō Sūgaku Danwakai*, **244** (1942), no. 1077, 1352–1400 (in Japanese).
7. Itô, K.: Stochastic integral, *Proc. Imp. Acad. Tokyo* **20** (1944), no. 8, 519–524.

8. Itô, K.: Multiple Wiener integral, *J. Math. Soc. Japan* **3** (1951) 157–169.
9. Itô, K.: Extension of stochastic integrals, *Proceedings of the International Symposium on Stochastic Differential Equations*, Kyoto, 1976, pp. 95–109, Wiley, New York-Chichester-Brisbane, 1978.
10. Khalifa, N., Kuo, H.-H., Ouerdiane, H., and Szozda, B.: Linear stochastic differential equations with anticipating initial conditions. *Communications on Stochastic Analysis* **7**, no. 2 (2013) 245–253.
11. Kubo, I. and Takenaka, S.: Calculus on Gaussian white noise III, *Proc. Japan Acad.* **57A** (1981) 433–437.
12. Kuo, H.-H.: *Gaussian Measures in Banach Spaces*, Lecture Notes in Math. **463**, Springer-Verlag, 1975.
13. Kuo, H.-H.: *White Noise Distribution Theory*, CRC Press, 1996.
14. Kuo, H.-H.: *Introduction to Stochastic Integration.* Universitext (UTX), Springer, 2006.
15. Kuo, H.-H., Peng, Y., and Szozda, B.: Itô formula and Girsanov theorem for anticipating stochastic integrals, *Communications on Stochastic Analysis* **7**, no. 3 (2013) 441–458.
16. Kuo, H.-H., Peng, Y., and Szozda, B.: Generalization of the anticipative Girsanov theorem, *Communications on Stochastic Analysis* **7**, no. 4 (2013) 573–589.
17. Kuo, H.-H., Sae-Tang, A., and Szozda, B.: A stochastic integral for adapted and instantly independent stochastic processes, in "*Advances in Statistics, Probability and Actuarial Science*" Vol. I, Stochastic Processes, Finance and Control: A Festschrift in Honour of Robert J. Elliott (eds.: Cohen, S., Madan, D., Siu, T. and Yang, H.), World Scientific, 2012, 53–71.
18. Kuo, H.-H., Sae-Tang, A., and Szozda, B.: The Itô formula for a new stochastic integral, *Communications on Stochastic Analysis* **6**, no. 4 (2012) 603–614.
19. Kuo, H.-H., Sae-Tang, A., and Szozda, B.: An isometry formula for a new stochastic integral, In "*Proceedings of International Conference on Quantum Probability and Related Topics*," May 29–June 4, 2011, Levico, Italy, *QP–PQ: Quantum Probability and White Noise Analysis* **29** (2013) 222–232.
20. Malliavin, P.: Stochastic calculus of variation and hypoelliptic operators, *Proceedings of the International Symposium on Stochastic Differential Equations*, Kyoto, 1976, pp. 195–263, Wiley, New York-Chichester-Brisbane, 1978.
21. Malliavin, P.: *Stochastic Analysis*, Springer-Verlag, Berlin 1997.
22. Obata, N.: *White Noise Calculus and Fock Space,* Lecture Notes in Math. **1577**, Springer-Verlag, 1994
23. Potthoff, J. and Streit, L.: A characterization of Hida distributions, *J. Funct. Anal.* **101** (1991) 212–229.
24. Skorokhod, A. V.: On a generalization of a stochastic integral, *Theory Probab. Appl.* **20** (1975) 219–23.
25. Wiener, N.: Differential space, *J. Math. Phys.* **58** (1923) 131–174.
26. Wiener, N.: The homogeneous chaos, *Amer. J. Math.* **60** (1938) 897–936.

Hui-Hsiung Kuo: Department of Mathematics, Louisiana State University, Baton Rouge, LA 70808, USA

E-mail address: kuo@math.lsu.edu