# Chapter 2
# Preliminary – Causal Calculus

The theory of noncausal calculus is an alternative to the causal theory of Itô calculus but is not quite independent of it. As we will see in the main part of this book that starts from Chap. 3, our noncausal theory stands as a natural extension of the causal theory of Itô calculus, to be more precise, the causal theory based on the stochastic integral called *symmetric integrals*. We may emphasize that at this point our noncausal theory keeps a large part of its *raisons d'être*.

Hence as preliminaries for the study of our noncausal theory of stochastic calculus we need to present in this chapter a necessary and minimum review on those materials and related facts from the causal calculus such as Brownian motion, the Itô integral, the symmetric integrals and the notion of the B-derivative of random functions. By doing this we also intend to prepare the list of symbols and terminologies concerning those materials that will be used throughout the book. We remark that what we intend to show in this chapter is not a standard review of Itô calculus but just a small note on it, thus for the details or further understanding of the causal calculus we would refer the reader to other standard textbooks on Itô calculus and some of the author's articles (e.g. [20–25]).

The presentation of these materials is in the following order: Brownian motion in Sect. 2.1, the Itô integral and related statements in Sect. 2.2, some elementary but important results concerning the SDE (stochastic differential equation) will be referred to in Sect. 2.3, while Sect. 2.4 is devoted to the note on variants of the Itô integral, where we repeat briefly the results concerning the B-derivative and the symmetric integrals, especially the integral $I_{1/2}(f)$ that is introduced by the author ([20, 21]) and will be of frequent use in the discussions on our main theme. We also refer to the integral of symmetric type called the Stratonovich–Fisk integral ([14, 55]).

Before entering into the discussion the author would like to have the reader's attention on the symbols for stochastic integrals. As we are going to deal with plural stochastic integrals we need appropriate symbols to make clear distinctions between them. In particular, for the Itô integral we would assign $\int f \, d_0 W_t$, by putting "0" at "$dW$" to signify that the Itô integral is at the origin of the theory of stochastic calculus.



S. Ogawa, *Noncausal Stochastic Calculus*, DOI 10.1007/978-4-431-56576-5_2

## 2.1 Brownian Motion

**Definition 2.1** (1) A real-valued random variable $X(\omega)$ defined on a probability space $(\Omega, \mathscr{F}, P)$ is called Gaussian if its characteristic function $\varphi_X(\theta) := E[\exp\{i\theta X\}]$ is given in the following form:

$$E[\exp\{i\theta X(\omega)\}] := \int_{\Omega} \exp\{i\theta X(\omega)\} dP(\omega) = \exp\left\{ im\theta - \frac{\sigma^2}{2}\theta^2 \right\} \quad {}^{\forall}\theta \in \mathbf{R}^1,$$

where $m$ and $\sigma \geq 0$ are real constants.
(2) An n-tuple of real random variables $\mathbf{X} := (X_1, \dots, X_n) \in \mathbf{R}^n$ is called an $\mathbf{R}^n$-valued (or $n$-dimensional) Gaussian random variable provided that any linear combination $Y = \sum_{k=1}^n t_k X_k(\omega)$ with ${}^{\forall}(t_1, \dots, t_n) \in \mathbf{R}^n$ is a real Gaussian random variable.

By definition (2) above we see that:

**Proposition 2.1** *The n components* $\{X_k, k = 1, n\}$ *of the* $\mathbf{R}^n$*-valued Gaussian variable* $\mathbf{X} = (X_1, \dots, X_n)$ *are independent provided that they are uncorrelated, namely,* $Cov(X_i, X_j) = 0 {}^{\forall} i \neq j$*, where* $Cov(X_i, X_j) = E[(X_i - EX_i)(X_j - EX_j)]$ *is the covariance of* $(X_i, X_j)$*.*

**Definition 2.2** (*Gaussian process*) A stochastic process $X_t(\omega),\ t \in \mathbf{T} \subset \mathbf{R}$ is called *Gaussian* provided that for any $n \in \mathbf{N}$ and arbitrarily chosen $n$ different points $t_i \in \mathbf{T},\ 1 = 1, \dots, n$, the $n$-dimensional random variable $(X_{t_i}(\omega), \dots, X_{t_n})$ is Gaussian.

We know that every finite dimensional Gaussian distribution is determined by a pair of parameters, namely, a mean vector $\mathbf{m} ={}^t (m_1, m_2, \dots, m_n) \in \mathbf{R}^n$ and an $n \times n$-real symmetric positive definite matrix $\Gamma$ called the covariance matrix, in the following way:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\Gamma|}} \exp\{-\frac{1}{2}(\Gamma^{-1}(\mathbf{x} - \mathbf{m}), \mathbf{x} - \mathbf{m})\}, \quad \mathbf{x} ={}^t (x_1, x_2, \dots, x_n) \in \mathbf{R}^n.$$

Thus we notice that a Gaussian process $X_t, t \in \mathbf{T}$ is completely determined by the pair of a mean function $m(t)$ and a real kernel $\Gamma(s,t),\ s, t \in \mathbf{T}$ of positive definite type. Notice also that a real Gaussian process $X_t$ determined by these has the following properties:

$$\begin{aligned} &m(t) := E[X_t], \\ &\Gamma(s,t) := Cov(X_s, X_t) = E[(X_s - m(s))(X_t - m(t))]. \end{aligned}$$

We introduce one of our principal materials, the Brownian motion (or the BM for short), in the following:

**Definition 2.3** (*Brownian motion*) (1) A real Gaussian process $W_{\cdot}(\omega)$ defined on $(\Omega, \mathscr{F}, P)$ is called Brownian motion provided that:

**(b1)** $P\{W_0 = 0\} = 1$,
**(b2)** $E[W_t] = 0,\ E[W_s W_t] = s \wedge t$ for ${}^\forall s, {}^\forall t \geq 0$ where $s \wedge t := \min\{s, t\}$.

(2) Let $W_1(t), W_2(t), \ldots, W_n(t)$ be $n$ independent copies of the Brownian motion. The $R^n$-valued Gaussian process $\mathbf{W}(t) =^t (W_1(t), W_2(t), \ldots, W_n(t))$ is called the $n$-dimensional Brownian motion.

*Example 2.1* The following processes $X_t$ are all Brownian motions, where $c$ is a positive constant: (1) $X_t = W_{t+c} - W_c$, (2) $X_t = \frac{1}{\sqrt{c}} W_{ct}$, (3) $X_t = t W_{1/t}$ $(t > 0)$ with convention $X_0 = 0$.

A right continuous and increasing family $\{\mathscr{F}_t, t \geq 0\}$ of sub $\sigma$-fields of $\mathscr{F}$ is called "filtration":

$$\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F} \quad {}^\forall s < t, \quad \text{and} \quad \mathscr{F}_t = \bigcap_{h>0} \mathscr{F}_{t+h}.$$

For instance, $\mathscr{G}_t^W := \sigma\{W_s \,|\, s \leq t\}$ or $\mathscr{G}_t^W \vee \sigma\{V\}$ where $V(\omega)$ is a random variable independent of Brownian motion are filtrations.

**Definition 2.4** In this book, by natural filtration we understand a right continuous and increasing family of sub $\sigma$-fields $\{\mathscr{F}_t^W,\ t > 0\}$ such that

$$\mathscr{F}_t^W \supset \mathscr{G}_t^W \quad {}^\forall t$$

and that for any $s \leq t$ increment $W_t - W_s$ is independent of $\mathscr{F}_s^W$. Here we understand that every sub $\sigma$-field $\mathscr{F}_t^W$ is completed with all $P$-null sets.

### *2.1.1 Some Properties of BM*

The Brownian motion process, which is also called the *Wiener process*, was introduced by N. Wiener in 1930. Being one of the most important materials in the theory of stochastic processes, it has been studied extensively by many authors, and many books have been published. We do not intend to repeat in detail, even some parts of, its basic properties but we shall content ourselves in this subsection with listing only some of its remarkable properties which cannot be missed for our present purpose:

(f1) From (b2) we see that $E[(W_t - W_s)^2] = |t - s|$ and that the random variable $W_t - W_s$ follows the normal law, $W_t - W_s \sim N(0, |t - s|)$.
(f2) The condition (b2) also implies that, for any $0 \leq s \leq t \leq u \leq v$, we have

$$\begin{aligned} E[(W_v - W_u)(W_t - W_s)] &= E[W_v W_t - W_t W_u - W_v W_s + W_u W_s] \\ &= t - t - s + s = 0. \end{aligned}$$

By virtue of Proposition 2.1 we see from (f2) that Brownian motion is a *process of independent increments*.

(f2)' Or in other words, for any $s \leq t$ the increment $\Delta_{s,t} W := W(t) - W(s)$ is independent of the field $\mathscr{G}_s^W$.
This property implies that Brownian motion is a martingale along with the family of sub $\sigma$-fields $\{\mathscr{G}_t^W\}_{t>0}$, that is, for any $t \geq s$ the following holds:

$$E[W_t|\mathscr{G}_s^W] = W_s \quad P - a.s.$$

(f3) Brownian motion $W_t$ is a martingale with respect to any natural filtration mentioned in Definition 2.4,

$$E[W_t|\mathscr{F}_s^W] = W_s \;\; P - a.s. \; {}^{\forall}t \geq s.$$

(f4) For any fixed $\alpha \in \mathbf{R}$ the process $Z_t = \exp\{\alpha W_t - \frac{\alpha^2 t}{2}\}$ becomes an $\mathscr{F}_t^W$-martingale. In fact, for any $t \geq s$ we have

$$\begin{aligned} E[Z_t|\mathscr{F}_s^W] &= E\left[Z_s \exp\left\{\alpha W_t - \frac{\alpha^2(t-s)}{2}\right\} |\mathscr{F}_s^W\right] \\ &= Z_s E\left[\exp\left\{\alpha(W_t - W_s) - \frac{\alpha^2(t-s)}{2}\right\} |\mathscr{F}_s^W\right] \\ &= Z_s \;\; P - a.s. \end{aligned}$$

since

$$E\left[\exp\left\{\alpha(W_t - W_s) - \frac{\alpha^2(t-s)}{2}\right\} |\mathscr{F}_s^W\right] = E\left[\exp\left\{\alpha W_{t-s} - \frac{\alpha^2(t-s)}{2}\right\}\right] = 1.$$

(f5) The fact (f2) also implies that Brownian motion is a homogeneous Markov process having the following kernel as transition probability density:

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\}, \quad x, y \in \mathbf{R}, \; t > 0. \tag{2.1}$$

Knowing the transition probability density we can construct the Markov process, so we confirm the existence of BM.

Here are some important properties concerning the regularity of the sample path of BM.
(f6) We notice the following property.

**Proposition 2.2** *Almost every sample path of BM is continuous but is not of bounded variation on any finite interval.*

For the verification of this statement we appeal to the following result called *Kolmogorov's test*, whose proof is omitted.

**Theorem 2.1** *If a real-valued stochastic process $X_t$ $t \in [0, T]$ satisfies the following condition for some positive constants $\alpha, \beta, C$:*

$$E[|X_t - X_s|^{\alpha}] \leq C|t - s|^{1+\beta} \quad \forall \ s, t \in [0, T], \tag{2.2}$$

*then almost every sample function of $X_t$ is continuous.*

*Proof* Now we verify the validity of Proposition 2.2. From the fact (f1) we have,

$$E[(W_t - W_s)^4] = 3(t - s)^2.$$

Hence we see the continuity of the sample function by virtue of Kolmogorov's test (Theorem 2.1) cited above.

For the verification of the second assertion, we put

$$V_n = \sum_{i=0}^{n} \left| W\left(\frac{i+1}{n}\right) - W\left(\frac{i}{n}\right) \right|.$$

It suffices to show that

$$\lim_{n\to\infty} V_n = \infty \ \ P - a.s.$$

Notice that the condition (b1) together with (f1) and (f2) implies the following inequality,

$$\begin{aligned} E[e^{-V_n}] &= \Pi_{i=0}^{n} E[e^{-|W(\frac{i+1}{n})-W(\frac{i}{n})|}] = \{E[e^{-|W(\frac{1}{n})|}]\}^n \\ &\leq E\left[1 - \left|W\left(\frac{1}{n}\right)\right| + \frac{1}{2}\left|W\left(\frac{1}{n}\right)\right|^2\right]^n \\ &\leq \left\{1 - \frac{1}{\sqrt{n}} + \frac{1}{2n}\right\}^n \longrightarrow 0, \ (\text{as } n \to \infty). \end{aligned}$$

Hence we see that $\lim_{n\to\infty} V_n = \infty$ almost surely and this implies the conclusion. □

The properties (f7), (f8) below concern the regularity of sample paths of Brownian motion. The proofs can be found in every standard textbook (cf. [15]) and are omitted here for the sake of making the content of this chapter as compact as possible.

(f7) Almost every sample path of the BM is not differentiable at almost every $t \in [0, T]$.

(f8) As for the modulus of continuity of $W$, we have the following result due to P. Lévy (cf. [15]):

$$P\left[\limsup_{\substack{0\le s\le t\le 1\\ h=t-s\downarrow 0}} \frac{|W(t)-W(s)|}{\sqrt{2h\log\frac{1}{h}}}=1\right]=1. \tag{2.3}$$

### 2.1.2 Construction of BM

We would like to finish this section with a note on the existence of the Brownian motion, since we could find there a basic idea that leads us to the noncausal stochastic integral. We have already mentioned in (f5) how the BM is constructed as a Markov process. Here we shall show different ways for the construction.

1. Construction by a Fourier series.
   Let $\{\varphi_n(t)\}$ be an orthonormal basis in $L^2(0,1)$ and let $\{\Xi_n(\omega)\}$ be an *i.i.d.* family of random variables following the standard normal law $N(0,1)$. Given these, consider a sequence $\{X_n(t,\omega)\}_n$ of random functions defined in the following way:

   $$X_n(t,\omega) := \sum_{k=1}^{n} \Xi_n(\omega)\int_0^t \varphi_k(s)ds.$$

   Notice that by Perseval's equality we have

   $$\sum_{k=1}^{\infty}\left|\int_0^t \varphi_k(s)ds\right|^2 = ||\mathbf{1}_{[0,t]}(\cdot)||_{L^2}^2 = t,$$

   and notice that this convergence is uniform in $t\in[0,1]$. Hence,

   $$\lim_{m,n\to\infty}\int_0^1 E\left[|X_n(t)-X_m(t)|^2\right]dt = \int_0^1 \sum_{k=m+1}^{n} |\int_0^t \varphi_k(s)ds|^2 dt = 0.$$

   In other words the sequence $\{X_n(t,\omega)\}_n$ converges in $L^2([0,1]\times\Omega)$ to a limit, say $X(t,\omega)$. We see that $E[X(t)] = \lim_n E[X_n(t)] = 0$ and that

   $$\begin{aligned}
   &\mathrm{Cov}(X(s),X(t)) = \lim_{n\to\infty}\mathrm{Cov}(X_n(s),X_n(t))\\
   &= \lim_{n\to\infty}\sum_{k=1}^{n}\int_0^s \varphi_{k(r)dr}\int_0^t \varphi_{k(r)dr} = (\mathbf{1}_{[0,s]}(\cdot),\mathbf{1}_{[0,t]}(\cdot))_{L^2}\\
   &= s\wedge t,
   \end{aligned}$$

   which shows that the limit $X$ is a Brownian motion.

2. As for the convergence of the series $X(\cdot)$ we must refer to a much more general result due to K. Itô and M. Nisio:

**Theorem 2.2** (Itô and Nisio [12]) *For an arbitrary orthonormal basis* $\{\varphi_n(t)\}$ *in* $L^2(0,1)$ *the series*

$$X(t.\omega) = \sum_n \Xi_n \tilde{\varphi_n}(t), \quad \text{where } \tilde{\varphi_n}(t) = \int_0^t \varphi_n(s)ds \tag{2.4}$$

*converges uniformly in $t$ over* $[0,1]$ *with probability one.*

For the proof we would refer the reader to the article [12] cited above.

3. Instead we would like to show the result due to Ciesielski [1] which deals with the series (2.4) for a special basis and can be verified in an elementary way:
Let $\{H_{n,i},\ 0 \le i \le 2^{n-1}-1,\ n \in \mathbf{N} \cup \{0\}\}$ be the orthonormal system of Haar functions, namely

$$\begin{aligned} &H_{0,0}(t) = 1, \quad t \in [0,1], \\ &H_{n,i}(t) = 2^{\frac{n-1}{2}}\{\mathbf{1}_{[2^{-n+1}i, 2^{-n+1}(i+1/2))}(t) - \mathbf{1}_{[2^{-n+1}(i+1/2), 2^{-n+1}(i+1))}(t)\}, \\ &\qquad n \ge 1,\ 0 \le i \le 2^{n-1}-1, \end{aligned} \tag{2.5}$$

where $\mathbf{1}_A(\cdot)$ is the indicator function of set $A$.
Given this we take a family of independent and identically distributed $N(0,1)$ random variables $\{\Xi_{0,0},\ \Xi_{n,i};\ 0 \le i \le 2^n-1,\ n \in \mathbf{N}\}$ and consider the random series as follows:

$$X(t) = \Xi_{0,0}t + \sum_{n=1}^{\infty}\sum_{i=0}^{2^n-1} \Xi_{n,i}\tilde{H}_{n,i}(t), \tag{2.6}$$

where

$$\tilde{H}_{n,i}(t) = \int_0^t H_{n,i}(s)ds, \quad t \in [0,1].$$

**Proposition 2.3** (Ciesielski [1]) *The series $X(t,\omega)$ converges uniformly in $t$ over* $[0,1]$ *with probability one and the sum $X$ is a Brownian motion:*

$$P\left[\lim_{m,n\to\infty} \sup_{t\in[0,1]} |\sum_{k=m}^{n}\sum_{i=0}^{2^k-1} \Xi_{k,i}\tilde{H}_{k,i}(t)| = 0\right] = 1.$$

*Proof Sketch of the proof*:
Put $X(t,\omega) := \Xi_{0,0}t + \sum_{n=1}^{\infty} Y_n(t)$, where

$$Y_n(t,\omega) = \sum_{i=0}^{2^n-1} \Xi_{n,i}\tilde{H}_{n,i}(t), \quad n \ge 1.$$

We are going to show that the series $\sum_{n=1}^{\infty} Y_n(t, \omega)$ converges uniformly in $t \in [0, 1]$ with probability one.

The functions $\tilde{H}_{n,k}(t) = \int_0^t H_{n.k}(s)ds$ are just the functions called the Schauder basis, each of which has an equi-lateral triangular shape with height $2^{-\frac{n+1}{2}}$. Also we notice that

$$\tilde{H}_{n,k}(t)\tilde{H}_{n,j}(t) = 0 \;\text{ whenever }\; k \neq j. \tag{2.7}$$

Then by the property (2.7) we have the following estimate:

$$\sup_{t\in[0,1]} |Y_n(t)| = \left| \sum_{i=0}^{2^n-1} \Xi_{n,i}(\omega)\tilde{H}_{n,i}(t) \right| \leq \max_i |\Xi_{n,i}| 2^{-(n+1)/2}.$$

Hence for an arbitrary positive $\alpha$ we get the inequality below:

$$\begin{aligned}
&P\{\sup_t |Y_n(t)| \geq \alpha\sqrt{2^{-n}\log 2^n}\} \\
&\leq P\{\max_i |\Xi_{n,i}| 2^{-(n+1)/2} \geq \alpha\sqrt{2^{-n}\log 2^n}\} \\
&\leq P\{\max_{0\leq i\leq 2^n-1} |\Xi_{n.i}| \geq \alpha\sqrt{2\log 2^n}\} \leq 2^n P\{|\Xi_{0,0}| \geq \alpha\sqrt{2\log 2^n}\} \\
&\leq 2^n 2 \int_{\alpha\sqrt{2n\log 2}}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx.
\end{aligned}$$

By the elementary inequality

$$\int_A^{\infty} e^{-x^2/2} \leq \frac{1}{A} e^{-A^2/2},$$

we get the following estimate,

$$P\{\sup_t |Y_n(t)| \geq \alpha\sqrt{2^{-n}\log 2^n}\} \leq \sqrt{\frac{2}{2\pi\log 2}} \cdot \frac{1}{\sqrt{n}} 2^{(1-\alpha^2)n}.$$

Since $\sum_n \frac{1}{\sqrt{n}} 2^{(1-\alpha^2)n} < \infty$ when $\alpha > 1$, we see by Borel–Cantelli's first lemma that

$$P\left[ \sup_{t\in[0,1]} |Y_n(t)| < \alpha\sqrt{2^{-n} n\log 2} \;\text{ for all large enough } n \right] = 1.$$

Since $\sum_n \sqrt{2^{-n} n \log 2} < \infty$, we confirm that the series $X(t, \omega)$ converges uniformly in $t \in [0, 1]$ with probability one. It is immediate to see that the process $X(t, \omega)$ is Gaussian and that.,

$$E[X(t)] = 0, \quad \mathrm{Cov}(X(s), X(t)) = s \wedge t,$$

hence we get the conclusion. □

## 2.2 Itô Integral with Respect to BM

Let $W_t(\omega)$ or $W(t, \omega)$, $(t \geq 1)$ be Brownian motion defined on a probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t^W\}, P)$, where $\{\mathscr{F}_t^W\}$ is the natural filtration mentioned in Definition 2.4. We review in this section how the Itô integral with respect to Brownian motion, the first stochastic integral, $\int f(t, \omega) d_0 W_t$ is introduced for a certain class of random functions.

### *2.2.1 Classes of Random Functions*

We need to introduce some classes of random functions. First of all, by random function we understand in this book a real or complex-valued function $f(t, \omega)$ which is defined on the complete measure spaces $(\mathbf{R}^1 \times \Omega, dt \times dP)$ and is measurable in $(t, \omega)$ with respect to the product $\sigma$-field $\mathscr{B}_{R_+} \times \mathscr{F}$, where $\mathscr{B}_{R_+}$ is the Borel field on $\mathbf{R}_+ = [0, \infty)$. For the simplicity of argument and notations we restrict ourselves to the case of random functions $f(t, \omega)$ defined on the unit interval $t \in [0, 1]$, but depending on the subject this restriction will be changed in a customary way to a case of random functions defined on a larger interval like functions on a finite interval $[0, T]$ or on $\mathbf{R}_+$.

Here is the list of symbols for classes of random functions which will be in frequent use throughout the book.

- **H**: The totality of such random functions $f(t, \omega)$ that verify the condition

$$P\left\{\int_0^1 |f(t, \omega)|^2 dt < \infty\right\} = 1.$$

- **M**: Set of all such random functions $f(t, \omega) \in \mathbf{H}$ that are adapted to the filtration $\{\mathscr{F}_t^W\}_{t>0}$ and, more precisely, are progressively measurable in $(t, \omega)$ with respect to the product field $\mathbf{B}_{[0,t]} \times \mathscr{F}_t^W$. We will call this constraint on the measurability of random functions the *causality condition* and call the random function of this class *causal*.
- **Note**: When we say that a random function $f(t, \omega)$ is *noncausal*, it means that the function is *not assured to be causal*, in other words it simply means that $f \in \mathbf{H}$. This may be an abuse of the word *noncausal*: nevertheless the word has been in use since the beginning of the theory, so also in this book we would like to follow this custom and hope that the reader will not be confused.

- $\mathbf{M}_2$: The set of all causal random functions $f(t,\omega)$ that satisfy the condition $E[\int_0^1 |f(t,\omega)|^2 dt] < \infty$, i.e. $\mathbf{M}_2 = \mathbf{M} \cap L^2([0,1] \times \Omega, dt \times dP)$.
- $\mathbf{M}_{2c}$: The subset of $\mathbf{M}_2$, consisting of all elements that are uniformly continuous in the mean sense, namely $\lim_{h\to 0} \sup_t E[|f(t+h) - f(t)|^2] = 0$.
- $\mathbf{S}_0$: The totality of such a random function $f(t,\omega)$ whose sample path is almost surely a step function in $t \in [0,1]$, that is, there exists a finite partition $\{0 = t_0 < t_1 < \cdots < t_n = 1\}$ of $[0,1]$ and random variables $\{f_i(\omega), 0 \le i \le n-1\}$ such that
$$f(t,\omega) = f_i(\omega), \quad t \in [t_i, t_{i+1}),\ i = 0, \ldots, n-1. \tag{2.8}$$
The random function of this class is called *simple*.
- $\mathbf{S}$: The set of all causal simple functions, that is, $\mathbf{S} = \mathbf{S}_0 \cap \mathbf{M}$. Being "simple and causal" is equivalent to the fact that each random variable $f_i(\omega)$ $(i = 0, \ldots, n-1)$ in the form (2.8) is measurable with respect to the $\sigma$-field $\mathscr{F}_{t_i}^W$.
- $\mathbf{S}_2$: The set of all simple and causal random functions which are square integrable in $(t,\omega)$, namely $\mathbf{S}_2 = \mathbf{S}_0 \cap \mathbf{M}_2$.

Itô's stochastic integral of a causal random function $f(t,\omega) \in \mathbf{M}$ with respect to Brownian motion is introduced step by step in the following way.

### 2.2.2 Itô Integral for $f \in \mathbf{S}$

Let $f(t,\omega)$ be an $\mathbf{S}$-class random function. By definition of the class, there exists a partition $\{0 = t_0 < t_1 < \cdots < t_n = 1\}$ of $[0,1]$ and a family of random variables $\{f_i(\omega), 0 \le i \le n-1\}$ such that

$$f(t,\omega) = \sum_{i=0}^{n-1} f_i(\omega) \mathbf{1}_{[t_i, t_{i+1})}(t), \quad t \in [0,1], \tag{2.9}$$

Notice that each $f_i(\omega) = f(t_i,\omega)$ is $\mathscr{F}_{t_i}^W$ measurable.

**Definition 2.5** For a causal simple function $f(t,\omega)$ of the form (2.9) we put

$$I(f) := \sum_{i=0}^{n-1} f_i(\omega) \Delta_i W, \quad \Delta_i W = W(t_{i+1}) - W(t_i).$$

We call $I(f)$ the Itô integral of $f$ $(\in \mathbf{S})$ with respect to Brownian motion and denote it by $\int_0^1 f(t,\omega)\, d_0 W_t$.

Here we notice that the representation form (2.9) of a simple function is not unique, indeed it can be represented along a different partition, but the above definition of the integral $I(f)$ for $f \in \mathbf{S}$ does not depend on those representation forms.

For a sub-interval $[a, b] \subset [0, 1]$, it is clear that the function $\mathbf{1}_{[a,b]}(\cdot) f(\cdot, \omega)$ belongs to the class $\mathbf{S}$. Hence the stochastic integral on the sub-interval $[a, b]$ is well-defined in the following form:

$$\int_a^b f(t.\omega) d_0 W_t := I(\mathbf{1}_{[a,b]}(\cdot) f(\cdot)).$$

We will denote the integral $I(\mathbf{1}_{[0,t]} f)$ also by $I_t(f)$.

The stochastic integral $I(f)$ defines an application from $\mathbf{S}$ to $L^0(\Omega, dP)$.

**Proposition 2.4** *The integral $I(f)$ (defined on $\mathbf{S}$) has the following properties:*

*(1) Linearity: The application $I(f)$ is linear, that is, for any functions $f, g \in \mathbf{S}$ and constants $\alpha, \beta$ the following equality holds:*

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g).$$

*(2) Isometry: For a causal and square integrable simple function, $f(t, \omega) \in \mathbf{S}_2$, we have $E[I(f)] = 0$ and*

$$E[|I(f)|^2] = E[\int_0^1 |f(t, \omega)|^2 dt] = ||f||^2_{L^2([0,1]\times\Omega)}.$$

*In other words, the Itô integral defines an isometry from $\mathbf{S}_2$ ($\subset L^2([0, 1] \times \Omega, dt \times dP)$) to $L^2(\Omega, dP)$.*

*Proof* Property (1) is evident. As for the second equality in (2), we have

$$\begin{aligned}
E[|I(f)|^2] &= E\left[\sum_{i>j} \{f_i \overline{f_j} + \overline{f_i} f_j\} \Delta_i W \Delta_j W + \sum_k |f_k|^2 (\Delta_k W)^2\right] \\
&= \sum_{i>j} E\left[\{f_i \overline{f_j} + \overline{f_i} f_j\} \Delta_j W \cdot E\{\Delta_i W | \mathscr{F}^W_{t_i}\}\right] + \sum_k E\left[|f_k|^2 E\{(\Delta_k W)^2 | \mathscr{F}^W_{t_k}\}\right] \\
&= \sum_k E[|f_k|^2](t_{k+1} - t_k) = E \int_0^1 |f(t, \omega)|^2 dt,
\end{aligned}$$

and this implies the conclusion. □

**Proposition 2.5** *Let $f(t, \omega)$ be a causal simple function. Then the stochastic process $I_t(f) := I(\mathbf{1}_{[0,t]}(\cdot) f(\cdot))$ has the following properties:*

*(a) Almost all sample functions of $I_t(f)$ are continuous in $t$.*
*(b) The process $I_t(f)$ is an $\mathscr{F}^W_t$- martingale.*
*(c) When $f \in \mathbf{S}$ is real, the process $Z_t := \exp\{I_t(f) - \frac{1}{2} \int_0^t f^2(s) ds\}$ is a continuous $\mathscr{F}^W_t$ martingale and $E[Z_t] = 1 \; \forall t$.*

*Proof* (a) Let $\Delta$ be a partition associated to the simple function $f(t, \omega)$. Fix a $t \in (0, 1]$ and denote by $[t_k, t_{k+1})$ the sub-interval that contains the $t$, then

$$I_t(f) = \sum_{i=0}^{k-1} f_i(\omega)\Delta_i W + f_k(\omega)\{W(t) - W(t_k)\}$$

which shows the continuity in $t$ of the path of process $I_t(f)$.

(b) Let $s, t \in [0, 1]$ be such that $t \geq t_i \geq t_j \geq s \geq t_{j-1}$. Then we have

$$\begin{aligned} I_t(f) &= \sum_{k=0}^{i-1} f_k \cdot \Delta_k W + f_i(W_t - W_{t_i}) \\ &= \sum_{k=0}^{j-2} f_k \cdot \Delta_k W + f_{j-1}(W_s - W_{t_{j-1}}) + f_{j-1}(W_{t_j} - W_s) \\ &+ \sum_{k=j}^{i} f_k \Delta_k W + f_i(W_t - W_{t_j}), \end{aligned}$$

Hence we get

$$\begin{aligned} E[I_t(f)|\mathscr{F}_s^W] &= \sum_{k=0}^{j-2} f_k \cdot \Delta_k W + f_{j-1}(W_s - W_{t_{j-1}}) \\ &+ E\left[ f_{j-1}(W_{t_j} - W_s) + \sum_{k=j}^{i} f_k \cdot \Delta_k W + f_i(W_t - W_{t_j}) \Big| \mathscr{F}_s^W \right] \\ &= I_s(f) + E\Big[ f_{j-1} \cdot E[W_{t_j} - W_s | \mathscr{F}_{j-1}^W] \\ &+ \sum_{k=j}^{i} f_k E[\Delta_k | \mathscr{F}_k^W] + f_i E[W_t - W_{t_j} | \mathscr{F}_j^W] \Big| \mathscr{F}_s^W \Big] \\ &= I_s(f) \quad P - a.s. \end{aligned}$$

(c) Let $s, t$ be such that $t \geq t_i \geq s \geq t_{i-1}$. We have

$$I_t(f) = I_s(f) + f_{i-1}(W_{t_i} - W_s) + f_i(W_t - W_{t_i}),$$

and

$$\begin{aligned} \int_0^t f^2(r)dr &= \int_0^s f^2(r)dr + \int_s^t f^2(r)dr \\ &= \int_0^s f^2(r)dr + f_{i-1}^2(t_i - s) + f_i^2(t - t_i). \end{aligned}$$

Combining these we get the following equality:

$$
\begin{aligned}
Z_t &= Z_s \\
&\times \exp\left\{ f_{i-1}(W_{t_i} - W_s) - \frac{1}{2} f_{i-1}^2 (t_i - s) + f_i (W_t - W_{t_i}) - \frac{1}{2} f_i^2 (t - t_i) \right\}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
E[Z_t | \mathscr{F}_s^W] = Z_s \times E \Bigg[ &\exp\{ f_{i-1}(W_{t_i} - W_s) - \frac{1}{2} f_{i-1}^2 (t_i - s) \} \\
\times &E[\exp\{ f_i (W_t - W_{t_i}) - \frac{1}{2} f_i^2 (t - t_i) \} | \mathscr{F}_{t_i}^W ] \big| \mathscr{F}_s^W \Bigg].
\end{aligned}
$$

Since

$$
E[\exp\{ f_i (W_t - W_{t_i}) - \frac{1}{2} f_i^2 (t - t_i) \} | \mathscr{F}_{t_i}^W ] = 1 \ \ P - a.s.
$$

we see that

$$
\begin{aligned}
E[Z_t | \mathscr{F}_s^W] &= Z_s E \left[ \exp\{ f_{i-1}(W_{t_i} - W_s) - \frac{1}{2} f_{i-1}^2 (t_i - s) | \mathscr{F}_s^W \right] \\
&= Z_s \ \ P - a.s.
\end{aligned}
$$

This completes the proof of (c). □

From property (c) in Proposition 2.5, we get the following result:

**Proposition 2.6** *Let $f$ be a real and causal simple function. Then for any positive constants $a, b$, we have the following inequality:*

$$
P \left[ \sup_t \{ I_t(f) - \frac{a}{2} \int_0^t f^2(s, \omega) ds \} > b \right] \le e^{-ab}. \tag{2.10}
$$

*Proof* For a real $f \in \mathbf{S}$, we put

$$
Z_t(f) := \exp\{ I_t(f) - \frac{1}{2} \int_0^t f^2(s, \omega) ds \}.
$$

By property (c) we know that $Z_t(af)$ is an $\mathscr{F}_t^W$-martingale. For the left-hand side of the inequality in (2.10) we have the following expression:

$$
\begin{aligned}
&P\left[\sup_t\{I_t(f) - \frac{a}{2}\int_0^t f^2(s,\omega)ds\} > b\right] \\
&= P\left[\sup_t\{I_t(af) - \frac{a^2}{2}\int_0^t f^2(s,\omega)ds\} > ab\right] \\
&= P\left[\sup_t Z_t(af) > e^{ab}\right]
\end{aligned}
$$

Hence by applying the submartingale inequality (Corollary 10.1 in Chap. 9) to the last term in the above inequality, we get

$$P\left[\sup_t\{I_t(f) - \frac{a}{2}\int_0^t f^2(s,\omega)ds\} > b\right] \leq e^{-ab}E[Z_1(af)] = e^{-ab}.$$

This completes the proof. □

### *2.2.3 Extension to* $f \in \mathbf{M}$

We have introduced the stochastic integral $I(f)$ for the causal simple functions $f \in \mathbf{S}$. We show that the domain of the integral is extended to the class $\mathbf{M}$; the integral $I(f)$ thus introduced for $f \in \mathbf{M}$ we call the Itô integral and denote by $\int f(t,\omega)d_0W_t$.

There are many ways to achieve this aim. Here we basically follow the argument by H. McKean [15] which makes use of the exponential martingale $Z_t(f)$ ($f \in \mathbf{S}$).

Let us begin with the following lemma, the verification of which may be found in every standard textbook on calculus and is left to the reader.

**Lemma 2.1** *Let $f(t)$ be a deterministic function which is square integrable over* $[0, 1]$. *We suppose that the function is extended over a larger interval in such way that; $f(t) = 0$ outside of* $[0, 1]$, *then we have the following equality:*

$$\lim_{h\to 0}\int_0^1 |f(t+h) - f(t)|^2 dt = 0.$$

The next statement plays a key rôle in the discussion. The proof will be given in the Appendices (see Chap. 10).

**Proposition 2.7** *The class* $\mathbf{S}$ *is dense in* $\mathbf{M}$, *that is: for any $f \in \mathbf{M}$ there exists a sequence $\{f_n\}$ in* $\mathbf{S}$ *such that*

$$\lim_{n\to\infty}\int_0^1 |f(t,\omega) - f_n(t,\omega)|^2 dt = 0 \qquad P\text{-a.s.}$$

Now given a causal function $f \in \mathbf{M}$ and a positive $\alpha > 1$, we consider a sequence of causal simple functions $\{f_n\} \in \mathbf{S}$ such that

$$P\left\{\int_0^1 |f(t,\omega) - f_n(t,\omega)|^2 dt \le \frac{1}{2n^{2\alpha}} \quad \text{for any large enough } n\right\} = 1. \tag{2.11}$$

Notice that the existence of such a sequence is assured by Proposition 2.7, and that the following equality holds:

$$P\left\{\int_0^1 |f_n(t,\omega) - f_{n-1}(t,\omega)|^2 dt \le \frac{2}{n^{2\alpha}} \quad \text{for all large enough } n\right\} = 1. \tag{2.12}$$

Let us consider the sequence of random functions $\{I_t(f_n)\}$ defined by the integrals $I_t(f_n) = \int_0^t f_n(s,\omega)d_0 W_s$, then we have the following statement.

**Proposition 2.8** *For the sequence* $\{f_n\} \in \mathbf{S}$ *satisfying the condition (2.11) the sequence* $\{I_t(f_n)\}$ *converges uniformly in* $t$ *over* $[0,1]$ *with probability one, that is,*

$$P\left[\lim_{m,n\to\infty} \sup_{t\in[0,1]} |I_t(f_n) - I_t(f_m)| = 0\right] = 1.$$

*Proof* We put $g_n = f_n - f_{n-1}$. For any fixed positive constants $a, b$, we have from Proposition 2.6 the following estimate:

$$P\left[\sup_t \{I_t(g_n) - \frac{a}{2}\int_0^t g_n^2(s,\omega)ds\} > b\right] \le e^{-ab}.$$

Now fix another constant $c > 1$ and choose $a, b$ in the above inequality in the following way:

$$a = n^\alpha \sqrt{\log n}, \quad b = \frac{c}{n^\alpha}\sqrt{\log n},$$

then we get the following estimate:

$$P\left[\sup_{t\in[0,1]} \{I_t(g_n) - \frac{a}{2}\int_0^t g_n(s)^2 ds\} > b\right] \le e^{-c\log n} = \frac{1}{n^c}.$$

Since $\sum_n \frac{1}{n^c} < \infty$, by Borel–Cantelli's first lemma we get the following equality:

$$P\left[\sup_{t\in[0,1]} \{I_t(g_n) - \frac{a}{2}\int_0^t g_n(s)^2 ds\} \le b \quad \text{for large enough } n\right] = 1,$$

which together with the estimate (2.12) implies that

$$\begin{aligned}1 &= P\left[\sup_{t\in[0,1]} I_t(g_n) \le b + \frac{a}{2}\int_0^1 g_n^2(s)ds \ \text{ for large enough } n\right]\\ &\le P\left[\sup_t I_t(g_n) \le \frac{2\sqrt{\log n}}{n^\alpha} \ \text{ for all large enough } n\right].\end{aligned}$$

Repeating the same argument with $-g_n$ instead of the $g_n$, we find that

$$P\left[\sup_t |I_t(g_n)| \le \frac{2\sqrt{\log n}}{n^\alpha} \ \text{ for all large enough } n\right] = 1.$$

Since $\sum_n \frac{\sqrt{\log n}}{n^\alpha} < \infty$ and since $I_t(g_n) = I_t(f_n) - I_t(f_{n-1})$, we confirm that the sequence $\{I_t(f_n)\}$ almost surely converges in uniform topology. Hence we are done. □

We remark at this stage that the $\lim_n I(f_n)$ does not depend on the choice of the approximating sequence $\{f_n\}(\in \mathbf{S})$, hence we arrive at the following definition of the Itô integral.

**Definition 2.6** (*Itô integral*) For the sequence $\{I(f_n)\}$ constructed in Proposition 2.8 we call the $\lim_n I(f_n)$ the Itô integral of the causal random function $f \in \mathbf{M}$ and denote it by $\int_0^1 f(t,\omega)d_0W_t$.

*Remark 2.1* (Integration over a general interval) It is only for the simplicity of the argument that we have limited our discussion to the integration over the interval [0, 1]. We clearly see that our argument works for the case of any finite interval $[0, T]$ $T < \infty$. Extension to the integration over the infinite interval $[0, \infty)$ can be carried out in a similar way.

For a causal random function $f(t,\omega)$ such that $\int_0^\infty |f(t,\omega)|^2 dt < \infty$ P-a.s. Since $\lim_{A\to\infty}\int_A^\infty |f(t,\omega)|^2 dt = 0$ (P-a.s.), we choose an increasing sequence of real numbers $\{c_n\}$ in the following way:

$$P\left\{\int_{c_n}^\infty |f(t,\omega)|^2 dt < \frac{1}{2n^{2\alpha}} \ \text{ for large enough } n\right\} = 1,$$

and we put $f_n(t,\omega) = \mathbf{1}_{[0,c_n]}(t) f(t,\omega)$. Then we see that for each $n$ the Itô integral $I(f_n) = \int_0^{c_n} f(t,\omega)d_0W_t$ is well-defined and that

$$P\left\{\int_0^\infty |f(t,\omega) - f_n(t,\omega)|^2 dt < \frac{1}{2n^{2\alpha}} \ \text{ for large enough } n\right\} = 1.$$

Now following a similar argument given in the proof of Proposition 2.8 we would confirm that the sequence $\{I(f_n)\}$ almost surely converges as $n$ tends to infinity. We define the stochastic integral of $f(t,\omega)$ over $[0,\infty)$ by its limit:

$$\int_0^\infty f(t,\omega)d_0W_t = \lim_{n\to\infty}\int_0^n f(t,\omega)d_0W_t.$$

Here are some basic properties of the Itô integral $I(f)$ ($f\in\mathbf{M}$), most of which can be verified by checking the limit procedure for the corresponding equalities given in Propositions 2.4 and 2.5.

**Proposition 2.9** *The stochastic integral $I(f)$ defined for $f\in\mathbf{M}$ has the following properties:*
*(I–1) Linearity: $I(f)$ defines a linear application from $\mathbf{M}$ to $L^0(\Omega, dP)$, that is, for any $f, g\in\mathbf{M}$ and any $\alpha,\beta\in\mathbf{C}$ the next relation holds:*

$$I(\alpha f+\beta g)=\alpha I(f)+\beta I(g).$$

*(I–2) Isometry: For $f\in\mathbf{M}_2$ we have $E[I(f)]=0$ and $||I(f)||^2_{L^2(dP)}=E[|I(f)|^2]=||f||^2_{L^2(dt\times dP)}$. But for the function $f\in\mathbf{M}$ we only have the inequality*

$$||I(f)||^2_{L^2(dP)}\le E[\int_0^1 f^2(t,\omega)dt]\ (\le\infty).$$

*(I–3) Continuity: Almost every sample of $I_t(f)=I(\mathbf{1}_{[0,t]}f)$ is continuous in $t$.*
*(I–4) Martingale property:*

*(i) For $f\in\mathbf{M}_2$ the process $I_t(f)$ is an $\mathscr{F}_t^W$-martingale.*
*(ii) Let $f(t,\omega)$ be in $\mathbf{M}$. Take an arbitrary positive number $A$ and let $\tau_A(\omega)=\max\{t>0,\ \int_0^t|f(s,\omega)|^2ds\le A\}$. Then $I_{\{t\wedge\tau\}}(f)$ is an $\mathscr{F}_t^W$-martingale.*

*Proof* The property (I–1) is evident from definition of the integral. For the verification of the first part of (I–2), we fix an $f(t,\omega)$ in $\mathbf{M}_2$ and take a sequence $\{f_n\}$ in $\mathbf{S}$ that converges to $f$ in $\mathbf{M}_2$, namely,

$$\lim_{n\to\infty}E\int_0^1|f(t,\omega)-f_n(t,\omega)|^2dt=0.$$

We know from (2) in Proposition 2.4 that

$$E[I(f_n)]=0\text{and}E[|I(f_n)|^2]=||f_n||^2_{L^2}([0,1]\times\Omega).$$

Letting $n\to\infty$ on both sides of these equalities we see that $E[I(f)]=0$, $E[|I(f)|^2]=||f||^2_{L^2([0,1]\times\Omega)}$.

Now suppose that $f\in\mathbf{M}$ and $\{f_n\}\in\mathbf{S}$ is such that

$$\lim_{n\to\infty}\int_0^1|f(t)-f_n(t)|^2dt=0\quad\text{P-a.s.}$$

Then we have

$$E[|I(f)|^2] = E[\liminf_{n\to\infty}|I(f_n)|^2] \le \liminf_{n\to\infty} E[|I(f_n)|^2]$$
$$= \lim_{n\to\infty} E[\int_0^1 |f_n|^2 dt] \le E[\lim_{n\to\infty}\int_0^1 |f_n|^2 dt] = E[\int_0^1 |f|^2 dt].$$

Hence we see the validity of property (I–2).

The property (I–3) being also immediate from the argument given in the proof of Proposition 2.8, we are going to verify the property (I–4).

(i) Let $f \in \mathbf{M}_2$, then we can choose a sequence $\{f_n\}$ in $\mathbf{S}_2$ such that $E[\int_0^1 |f(t,\omega) - f_n(t,\omega)|^2 dt] = 0$. For instance we may take

$$f_n(t,\omega) = f(\eta_n(t),\omega) \;\text{ where }\; \eta_n(t) = \frac{k}{n} \;\text{ for }\; t \in \left[\frac{k}{n}, \frac{k+1}{n}\right).$$

For each $f_n$ the $I_t(f_n)$ being an $\mathscr{F}_t^W$-martingale, we know that for any $t \ge s \ge 0$, the following holds:

$$E[I_t(f_n)|\mathscr{F}_s^W] = I_s(f_n) \quad \text{P-a.s.} \tag{2.13}$$

Fix an $\mathscr{F}_s^W$-measurable random variable $X$ in an arbitrary way, then we have from (2.13) the equality

$$E[X \cdot I_t(f_n)] = E[X \cdot E\{I_t(f_n)|\mathscr{F}_s^W\}] = E[X \cdot I_s(f_n)].$$

Letting $n \to \infty$ on both sides of the above equality, we get the equality $E[X \cdot I_t(f)] = E[X \cdot I_s(f)]$, which states that the $I_t(f)$ is an $\mathscr{F}_t^W$-martingale.

(ii) Notice that $\tau_A(\omega)$ is an $\mathscr{F}_t^W$-stopping time and that the function $\mathbf{1}_{\{t\le\tau_A\}}(\omega)f(t,\omega)$ belongs to the class $\mathbf{M}_2$. Since $I_{t\wedge\tau_A}(f) = I_t(\mathbf{1}_{\{\cdot\,\le\tau_A\}}\cdot f)$ we confirm the validity of the statement.

□

### 2.2.4 Linearity in Strong Sense

We have seen how the Itô integral is introduced for causal functions in **H** and we have listed some of its basic properties. In the following sections we are going to show how the stochastic calculus based on the Itô integral works. Before that we like to close this section with the note on a remarkable character of the Itô integral, which is often missed in standard textbooks.

- It is said that the Itô integral is a *Riemann-type integral* in the sense that the Itô integral can be defined as a limit of Riemann sums as follows:

Given a finite partition $\Delta = \{0 = t_0 < t_1 < t_2 < \cdots < t_i < \cdots t_n = 1\}$ on the interval, that is [0, 1] in our present set-up, we consider the sum

$$R_\Delta(f) := \sum_{t_i \in \Delta} f(t_i, \omega)\Delta_i W, \quad \text{where } \Delta_i W = W(t_{i+1}) - W(t_i).$$

and say that $\int_0^1 f(t,\omega)d_0W_t = \lim_{|\Delta|\to 0} R_\Delta(f)$, where $|\Delta| = \max\{t_{i+1} - t_i;\ t_i \in \Delta\}$.

By looking at the proof of Proposition 2.7 we are sure that this is true at least for any $f \in \mathbf{M}_c$. Remember that any $f \in \mathbf{M}_c$ is approximated by the sequence of causal simple functions $f_\Delta$ as follows:

$$f_\Delta(t,\omega) = f(\eta_\Delta(t),\omega), \quad \text{where } \eta_\Delta(t) = \sum_k t_k \mathbf{1}_{[t_k,t_{k+1})}(t).$$

We are also sure that

$$\int_0^1 f(t,\omega)d_0W_t = \lim_{|\Delta|\to 0} R_\Delta(f) \quad \text{in the mean}$$

when $f \in \mathbf{M}_{2,c}$, i.e. when $f$ is causal and continuous in the mean,

$$\lim_{h\to 0} \sup_{t\in[0,1]} E|f(t+h) - f(t)|^2 = 0.$$

In fact, we have

$$I(f) - R_\Delta(f) = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \{f(t) - f(t_i)\}d_0W_t,$$

hence

$$\lim_{|\Delta|\to 0} E[|I(f) - R_\Delta(f)|^2] = \lim_{|\Delta|\to 0} \sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} E[|f(t) - f(t_i)|^2]dt = 0.$$

Itô calculus was introduced for the theory of SDEs (stochastic differential equations) whose solutions in most cases belong to those classes cited above, hence the slogan saying "the Itô integral is a Riemann integral" has caused no problem.

- Anyhow, the fact that the Itô integral $I(f)$ for $f \in \mathbf{M}$ is defined along the sequence of approximate causal simple functions is important. Let $f(t,\omega)$ be a function in $\mathbf{M}$ and let $\alpha(\omega)$ be an arbitrary random variable. Then we see that the function $g(t,\omega) := \alpha(\omega)f(t,\omega)$ is no more causal and is excluded from the class $\mathbf{M}$. But it does not prevent us from defining the Itô integral of a noncausal function of this particular form. In fact, let $\{f_n\}$ be a sequence of causal simple functions that converges to $f(t, \omega)$, and let $g_n(t, \omega) := \alpha(\omega) f_n(t, \omega)$. Then $g_n(t, \omega)$ is simple but not causal, but still the sequence $\{g_n\}$ converges to $g(t, \omega)$ in $\mathbf{H}$ P-a.s. By the definition of the Itô integral for simple functions we have

$$I(g_n) = \alpha(\omega) I(f_n),$$

hence we see that $\{I(g_n)\}$ converges to $\alpha(\omega) I(f)$, namely we confirm that $g(t, \omega)$ is Itô integrable.

- Generally speaking, once an integral is defined for functions in some subspace $\mathscr{D}$ of a function space, it defines at the same time an application from that domain $\mathscr{D}$ to a set in a metric space. In the case of the Itô integral the application $I(f)$ is a mapping from $\mathbf{M}$ to $L^0(\Omega, dP)$. On the other hand, in analysis it is widely believed that any application induced by an integral is *linear*. This is true for all stochastic integrals that we treat in this book, but the mappings, say $T(f)$ for the moment, induced by the Itô integral or by the *noncausal integral of the author* (see Chap. 3) exhibit more strong linearity as follows:

$$T(\alpha(\omega) f + \beta(\omega) g) = \alpha(\omega) T(f) + \beta(\omega) T(g),$$

for any random variables $\alpha(\omega)$, $\beta(\omega)$. This property we like to call the *strong linearity*.

### *2.2.5 Itô Formula*

Let $X_t$ be a stochastic process that accepts the following representation:

$$X_t = \xi(\omega) + \int_0^t b(s, \omega) ds + \int_0^t a(s, \omega) d_0 W_s \tag{2.14}$$

where $\xi(\omega)$ is a random variable independent of the Brownian motion $W.$, and $a(\cdot)$, $b(\cdot)$ are causal functions.

**Definition 2.7** Every causal stochastic process $X_t$ of the form (2.14) is called an Itô process and the totality of all such processes will be denoted by $\mathbf{M}_I$. In other words the Itô process is a special type of the Brownian semi-martingale.

The expression (2.14) is also denoted by the following differential form:

$$dX_t = b(t, \omega) dt + a(t, \omega) d_0 W_t, \quad X_0 = \xi(\omega).$$

*Remark 2.2* As a variant of this, the process of the following form (2.15) is called a quasi martingale:

$$dX_t = dB(t, \omega) + a(t, \omega) d_0 W_t, \tag{2.15}$$

when $a(t, \omega) \in \mathbf{M}$ and $b(t, \omega)$ is a process, causal or not, almost every sample of which is of bounded variation on $[0, 1]$. We will denote the totality of quasi-martingales by $\mathbf{M}_Q$. The semi-martingale is the name for a causal quasi-martingale.

Let $X_t$ be an Itô process and let $F(t, x)$ be a real-valued function that is defined on $\mathbf{R}_+ \times \mathbf{R}^1$ and is of $C^1$-class in $t$ and of $C^3$-class in $x$ with bounded derivatives.

**Theorem 2.3** (Itô formula) *The stochastic process $F(t, X_t)$ satisfies the following equality:*

$$
\begin{aligned}
&dF(t, X_t)\\
&= \left\{ F'_t(t, X_t) + F'_x(t, X_t) b(t, \omega) + \frac{1}{2} F''_{xx}(t, X_t) b^2(t, \omega) \right\} dt + F'_x(t, X_t) a(t, \omega) d_0 W_t,\\
&\text{where } F'_t = \partial_t F,\ F'_x = \partial_x F,\ F''_{xx} = \partial_x^2 F. \qquad (2.16)
\end{aligned}
$$

*This equality is called the Itô formula.*

*Proof* Fix $t$ and set $t_k^n = \frac{k}{n} t, \quad k = 0, \ldots, n.$

Then by the mean value theorem in calculus we have the following equality:

$$
\begin{aligned}
F(t, X_t) - F(0, X_0) &= \sum_{k=0}^{n-1} \{ F(t_{k+1}^n, X(t_{k+1}^n)) - F(t_k^n, X(t_k^n)) \}\\
&= \sum_{k=0}^{n-1} \{ F(t_{k+1}^n, X(t_{k+1}^n)) - F(t_k^n, X(t_{k+1}^n)) + F(t_k^n, X(t_{k+1}^n)) - F(t_k^n, X(t_k^n)) \}\\
&= \sum_{k=0}^{n-1} \{ F'_t(t_k^n + \theta_k^1 \Delta_k^n, X(t_{k+1}^n)) \Delta_k^n + F'_x(t_k^n, X(t_k^n)) \Delta_k^n X\\
&\qquad + \frac{1}{2} F''_{xx}(t_k^n, X(t_k^n) + \theta_k^2 \Delta_k^n X)(\Delta_k^n X)^2 \},
\end{aligned}
$$

where $\theta_k^1, \theta_k^2$ $(k = 0, \ldots, n-1)$ are constants in $(0, 1)$ and

$$
\Delta_k^n = t_{k+1}^n - t_k^n,\ \ \Delta_k^n X = X(t_{k+1}^n) - X(t_k^n).
$$

By virtue of the smoothness of $F'_t$, $F'_x$ and the continuity of the sample of $X(t, \omega)$, we see that

$$
\lim_{n \to \infty} \sum_{k=0}^{n-1} F'_t(t_k^n + \theta_k^1 \Delta_k^n, X(t_{k+1}^n)) \Delta_k^n = \int_0^t F'_t(s, X(s)) ds, \quad \text{P-a.s.}
$$

and

$$\lim_{n\to\infty}\sum_{k=0}^{n-1}\{F'_x(t_k^n, X(t_k^n))\Delta_k^n X + \frac{1}{2}F''_{xx}(t_k^n, X(t_k^n) + \theta_k^2\Delta_k^n X)(\Delta_k^n X)^2\}$$
$$= \int_0^t F'_x(s, X(s))\left\{b(s,\omega)ds + a(s, X(s))d_0W_s + \frac{1}{2}F''_{xx}(s, X(s))b^2(s)ds\right\},$$

P-a.s.

This completes the proof. □

We know that Itô formula (2.16) is valid for a function $F(t, x)$ with a less restrictive condition on regularity, but we do not enter into the details since we need not do so for our discussion in later chapters.

*Example 2.2* Applying the formula (2.16) to the case where

$$F(x) = e^x, \quad X_t = \int_0^t f(s,\omega)d_0W_t - \frac{1}{2}\int_0^t f^2(s,\omega)ds,$$

we get the next equality for $Z_t = \exp\{X_t\}$,

$$dZ_t = Z_t\left\{f\cdot d_0W_t - \frac{1}{2}f^2dt\right\} + \frac{1}{2}Z_tf^2dt = Z_tf\ d_0W_t$$

from which we see that $Z_t$ is an $\mathscr{F}_t^W$-martingale.

The extension of the result (Theorem 2.3) to the case of multi-dimensional Itô process can be given as follows.

Let $\mathbf{X}_t =^t (X_t^1, X_t^2, \ldots, X_t^p)$ be a $p$-dimensional stochastic process, each component $X_t^i$ of which is generated by the following rule:

$$dX_t^i = b^i(t,\omega)dt + \sum_{k=1}^q a_k^i(t,\omega)d_0W_t^k, \quad 1 \le i \le p, \tag{2.17}$$

where $a_j^i$, $b^i(1 \le i \le p,\ 1 \le j \le q)$ are causal random functions and $\mathbf{W}_t =^t (W^1, W^2, \ldots, W^q)$ is the $q$-dimensional Brownian motion, namely the $\{W^1, W^2, \ldots, W^q\}$ are independent Brownian motions. The Itô formula is extended in the following way.

**Theorem 2.4** (Itô formula 2) *Let $F(t,\mathbf{x})$ $(t \in \mathbf{R}_+,\ \mathbf{x} =^t (x_1, x_2, \ldots, x_p) \in \mathbf{R}^p)$ be a smooth function that is of $C^1$-class in $t$ and of $C^3$-class in $\mathbf{x}$ with bounded derivatives. Then for the $p$-dimensional Itô process $\mathbf{X}_t$ given in (2.17) above we have the following equality:*

$$dF(t,\mathbf{X}_t) = L_1F(t,\mathbf{X}_t)dt + L_2F(t, X_t)d_0\mathbf{W}_t, \tag{2.18}$$

*where*

$$L_1 = \frac{\partial}{\partial t} + \sum_{i=1}^{p} b^i(t,\omega)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{p}\sum_{k=1}^{m} a_k^i a_k^j \frac{\partial^2}{\partial x_i \partial x_j},$$

*and*

$$L_2 F(t,\mathbf{X})d_0\mathbf{W}_t = \sum_{k=1}^{m}\sum_{i=1}^{p} b_i^k(t,\omega)\frac{\partial}{\partial x_i}F(t,\mathbf{X})d_0W_t^k.$$

The demonstration of this statement is left to the reader.

*Example 2.3* Apply the Itô formula (2.18) to the case where $F(x,y) = xy$, $X_t = \int_0^t f(s,\omega)d_0W_s$ and $Y_t = \int_0^t g(s,\omega)d_0W_s$ where $f,\ g \in \mathbf{M}$, then we find the following equality:

$$\begin{aligned}
&\int_0^t f(s,\omega)d_0W_s \int_0^t g(s,\omega)d_0W_s \\
&= \int_0^t f(s,\omega)d_0W_s \int_0^s g(r.\omega)d_0W_r + \int_0^t g(s,\omega)d_0W_s \int_0^s f(r,\omega)d_0W_r \\
&\quad + \int_0^t f(s,\omega)g(s.\omega)ds.
\end{aligned} \tag{2.19}$$

### 2.2.6 About the Martingale $Z_t(\alpha)$

Because of the remarkable property of the causal function

$$Z_t = \exp\left\{I_t(f) - \frac{1}{2}\int_0^t f^2(s,\omega)ds\right\} \quad (f \in \mathbf{M}),$$

we try to look at it from a different viewpoint.

Let $f \in \mathbf{M}$ be a real causal function such that

$$E\left[\exp\left\{\alpha^2\int_0^1 f^2 dt\right\}\right] < \infty \quad {}^{\forall}\alpha \in R,$$

and consider again the process $Z_t(\alpha)$ introduced in Example 2.2:

$$Z_t(\alpha) := \exp\left\{\alpha\int_0^t f(s,\omega)d_0W_s - \frac{\alpha^2}{2}\int_0^t f^2(s,\omega)ds\right\}.$$

We have seen that it is a square integrable $\mathscr{F}_t^W$-martingale such that

$$dZ_t(\alpha) = \alpha f(t,\omega)Z_t(\alpha)d_0W_t, \qquad Z_0(\alpha) = 1. \tag{2.20}$$

In other words,

$$Z_t(\alpha) = 1 + \alpha \int_0^t f(s,\omega) Z_s(\alpha) d_0 W_s,$$

by which we find the following equality:

$$\begin{aligned} Z_t(\alpha) &= 1 + \alpha \int_0^t f(s,\omega) \left\{ 1 + \alpha \int_0^s f(r,\omega) Z_r(\alpha) d_0 W_r \right\} d_0 W_s, \\ &= 1 + \alpha \int_0^t f(s_1,\omega) d_0 W_{s_1} + \alpha^2 \int_0^t f(s_1,\omega) d_0 W_{s_1} \left( \int_0^{s_1} f(s_2,\omega) d_0 W_{s_2} \right). \end{aligned}$$

By the formal repetitin of this procedure we would find the expression for $Z_t(\alpha)$ as follows;

$$\begin{aligned} &Z_t(\alpha) = \sum_{n=0}^{\infty} \alpha^n Z_n(t), \text{ where } Z_0(t) = 1 \text{ and} \\ &Z_n(t) = \int_0^t f(s_1,\omega) d_0 W_{s_1} \int_0^{s_1} f(s_2,\omega) d_0 W_{s_2} \ldots \int_0^{s_{n-1}} f(s_n,\omega) d_0 W_{s_n}. \end{aligned} \tag{2.21}$$

In fact we can show the following statement:

**Lemma 2.2** *For any $\alpha$ the series $\sum_{n=0} \alpha^n Z_n(t)$ in (2.21) almost surely converges to $Z_t(\alpha)$, uniformly in $t$ on any compact interval.*

To keep the size of this chapter compact we will give the proof of this statement in Chap. 10.

Let $\{h_n(t,x),\ t > 0,\ x \in \mathbf{R}\}$ be the family of Hermite polynomials each element of which is defined by the following formula:

$$h_n(t,x) := \frac{(-1)^n}{n!} \exp\left\{ \frac{x^2}{2t} \right\} \frac{\partial^n}{\partial x^n} \exp\left\{ -\frac{x^2}{2t} \right\} \qquad (n \geq 0).$$

By this definition we see that the family of Hermite polynomials $\{h_n(t,x)\}$ has $\exp\{\alpha x - \frac{\alpha^2 t}{2}\}$ as its generating function, namely

$$\exp\left\{ \alpha x - \frac{\alpha^2 t}{2} \right\} = \sum_{n=0}^{\infty} \alpha^n h_n(t,x). \tag{2.22}$$

Now by substituting $\int_0^t f(s,\omega) d_0 W_s$ and $\tau(t) = \int_0^t f^2(s,\omega) ds$ for $x$ and $t$ respectively in the equality (2.22) we get the following equality:

$$
\begin{aligned}
Z_t(\alpha) &= \exp\left\{\alpha \int_0^t f(s,\omega)d_0W_s - \frac{\alpha^2}{2}\int_0^t f^2(s,\omega)ds\right\} \\
&= \sum_{n=0}^{\infty} \alpha^n h_n\left(\tau(t), \int_0^t f(s,\omega)d_0W_s\right).
\end{aligned}
$$

Comparing this with the equality (2.21), we find the following expressions:

**Lemma 2.3** *For a real* $f \in \mathbf{M}$ *it holds that*

$$
\begin{aligned}
Z_n(t) &= \int_0^t f(s_1,\omega)d_0W_{s_1}\int_0^{s_1} f(s_2,\omega)d_0W_{s_2}\dots\int_0^{s_{n-1}} f(s_n,\omega)d_0W_{s_n} \\
&= h_n\left(\tau(t), \int_0^t f(s,\omega)d_0W_s\right).
\end{aligned}
$$

Hermite polynomial $h_n(t,x)$ is a polynomial of $x$ of degree $n$ which is explicitly computed by its definition. Therefore we find the formula mentioned in Lemma 2.3 useful to compute or estimate the $n$-th moment of the Itô integral $E\left[\left(\int_0^t f(s,\omega)d_0W_s\right)^n\right]$. For instance:

**Lemma 2.4** *For an* $f \in \mathbf{M} \cap L^4$ *the following estimate holds:*

$$
E\left[\left(\int_0^t f(s,\omega)d_0W_s\right)^4\right] \le 36E\left[\left(\int_0^t f^2(s,\omega)ds\right)^2\right].
$$

*Proof* We know that $h_4(t,x) = x^4 - 6tx^2 + 3t^2$. Since $E[Z_4(t)] = 0$ we have $E[h_4(\tau(t), \int_0^t f(s,\omega)d_0W_s)] = 0$, hence we find that

$$
\begin{aligned}
E[(\int_0^t fd_0W_s)^4] &= E\left[6\int_0^t f^2 ds\left(\int_0^t fd_0W_s)^2 - 3(\int_0^t f^2(s,\omega)ds\right)^2\right] \\
&\le 6E\left[\left(\int_0^t fd_0W_s\right)^2\int_0^t f^2(s,\omega)ds\right] \\
&\le 6\left(E\left[\left(\int_0^t fd_0W_s\right)^4\right]\right)^{1/2}\left(E\left[\left(\int_0^t f^2 ds\right)^2\right]\right)^{1/2}.
\end{aligned}
$$

From this we find the inequality. □

## 2.3 Causal Variants of Itô Integral

### *2.3.1 Symmetric Integrals*

As we have noticed in the preceding section, the Itô integral has such a remarkable property that the process defined by the Itô integral $X_t = \int^t f(s,\omega)d_o W_s$ becomes a martingale and this fact is granted by the causality condition on the integrand $f(t,\omega)$ and by the particular form of Riemann sum. Hence any change in these situations would cause the loss of that nice property. Nevertheless we may think of the stochastic integral for the causal function by a Riemann sum of different form as follows:

$$R^\theta_\Delta(f) := \sum_{i=0}^{n-1} f(t_i + \theta\Delta_i)\Delta_i W, \tag{2.23}$$

where $\theta$ is a constant such that $\theta \in [0, 1]$. We are also interested in the following Riemann sums:

$$S^\theta_\Delta(f) := \sum_{i=0}^{n-1}\{f(t_i) + \theta\Delta_i f\}\Delta_i W \quad \Delta_i f = f(t_{i+1},\omega) - f(t_i,\omega). \tag{2.24}$$

We may see that the Riemann sum $S^{1/2}_\Delta(f)$ is just a stochastic variant of the trapezoidal formula in classic calculus. We are going to study the convergence of these Riemann sums. For this purpose we need to introduce a kind of regularity of the random function $f(t,\omega)$ with respect to the Brownian motion.

Let $\{\Delta_n\}$ be a sequence of partitions $\Delta_n = \{0 = t^n_0 < \cdots < t^n_s = 1\}$ of $[0, 1]$. We call the sequence *regular* provided that

$$\Delta_n \subset \Delta_{n+1} \text{ and } \lim_{n\to\infty} |\Delta_n| = 0.$$

We say that a random function $h(t,\omega)$ is *B-negligible* if it exhibits the following property (W):

**(W)** For each fixed $t \in [0, 1]$ and for any regular sequence of partitions $\{\Delta_n\}$ it holds that

$$\lim_{n\to\infty} \sum_{t_i\in\Delta_n(t)} (\Delta_i h)\cdot(\Delta_i W) = 0 \ \ (in\ P), \ \text{ where } \Delta_n(t) := \Delta_n \cup \{t\}.$$

In particular we call $h(t,\omega)$ *strongly B-negligible* if it satisfies the following condition (WS):

**(WS)** For each fixed $t \in [0, 1]$ and for any regular sequence of partitions $\{\Delta_n\}$ it holds that

$$\lim_{n\to\infty} \sum_{t_i \in \Delta_n(t)} |\Delta_i h \cdot \Delta_i W| = 0 \ \ (in\ P), \quad \text{where} \ \ \Delta_n(t) := \Delta_n \cup \{t\}.$$

*Example 2.4* A random function $h(t, \omega)$ that satisfies the following condition, for any regular sequence of partitions $\{\Delta_n\}$ of [0, 1], is strongly B-negligible:

$$\lim_{n\to 0} \sum_{t_i \in \Delta_n} |\Delta_i h|^2 = 0 \quad \text{in P.}$$

Thus any function, almost every sample of which is of bounded variation, is strongly B-negligible.

**Definition 2.8** (*B-derivative*) A random function $f(t, \omega) \in \mathbf{H}$ is called *B-differentiable* (or *strongly B-differentiable*) if there exists a causal function $g(t, \omega) \in \mathbf{M}$ such that the function $h(t, \omega) = f(t, \omega) - \int_0^t g(s, \omega) d_0 W_s$ is B-negligible (or strongly B-neligible respectively). In this case the function $g(t, \omega)$ we call the B-derivative of $f(t, \omega)$ and denote it by the symbol $\hat{f}(t, \omega)$ or by $\frac{\partial}{\partial W_t} f(t, \omega)$.

We see the uniqueness of the B-derivative in the following statement whose proof will be given in the last chapter "Appendices 2":

**Proposition 2.10** *The B-derivative of a B-differentiable function is uniquely determined.*

*Remark 2.3* (B-differentiability [21, 43]) The notion of B-differentiability was first introduced in the study of the symmetric integral $\mathrm{I}_{1/2}$ and BPE (cf. [20–22]) where all random functions $f(t, \omega)$ are supposed to be causal. That was given in the following way.

A causal random function $f \in \mathbf{M}$ is called B-differentiable (or differentiable with respect to Brownian motion) if there exists a causal random function, say $\hat{f}(t, \omega)$, that satisfies the following condition:

$$\lim_{h\downarrow 0} \sup_{0\le t\le 1-h} \frac{1}{h} E\left[\left| f(t+h) - f(t) - \int_t^{t+h} \hat{f}(s) d_0 W_s \right|^2\right] = 0.$$

It may be immediate to see that new definition of B-differentiablity is a refinement of this classic one.

*Example 2.5* Let $X_t$ be a random function of the form

$$X_t = B(t, \omega) + \int_0^t a(s, \omega) d_0 W_s,$$

where $a(t, \omega)$ is a causal function of the class $\mathbf{M}_2$ and $B(t, \omega)$ a function which is Hölder continuous in $L^2(\Omega)$ sense, $E[|B(t+h) - B(t)|^2] = O(|h|^\alpha)$ $(1 < \alpha)$. Then $X$ is B-differentiable and $\frac{\partial X}{\partial W_t} = a(t, \omega)$.

In particular, every Itô process $X_t; dX_t = b(t,\omega)dt + a(t.\omega)d_0 W_t$ with $a(t,\omega) \in \mathbf{M}$, is B-differentiable.

After the introduction of the notion of B-differentiablity we find it convenient to denote by $\mathbf{H}^1$ the totality of all B-differentiable functions causal or not, and by $\mathbf{M}^1$ the set of all B-differentiable causal functions, i.e. $\mathbf{M}^1 := \mathbf{H}^1 \cap \mathbf{M}$.

We have the following result whose proof is given in Chap. 10 (Appendices–2);

**Theorem 2.5** ([21]) *Let $\{\Delta_n\}$ be an arbitrary sequence of partitions in $[0,1]$ such that, $\Delta_n \subset \Delta_{n+1}$ and $\lim_{n\to\infty} ||\Delta_n|| = 0$. Then for every B-differentiable function $f \in \mathbf{M}^1$ and a fixed $\theta \in [0,1]$, the two sequences of Riemann sums $R^\theta_{\Delta_n}(f)$, $S^\theta_{\Delta_n}(f)$ converge in probability to the same limit $I_\theta(f)$ which we also denote by $\int_0^1 f\, d_\theta W_t$ and call the $\theta$-integral. The integral $I_\theta(f)$ is related to the Itô integral $I_0(f)$ in the following form;*

$$I_\theta(f) = I_0(f) + \theta \int_0^1 \hat{f}(t,\omega)dt, \quad \hat{f} = \frac{\partial f}{\partial W_t}.$$

Among these integrals $I_\theta$, the two $I_0$ and $I_{1/2}$ are of particular importance, the former is of course the Itô integral and the latter we call the *symmetric integral*.

*Example 2.6* (a formula concerning the white noise) Let $X_t$ be a causal function defined by the symmetric integral as follows:

$$X_t = X_0 + \int_0^t f(s,\omega)dW_s, \quad f(t,\omega) \in \mathbf{M}^1$$

for some B-differentiable function $f(t,\omega)$.

Notice that we can verify the validity of the following expression in the sense of L. Schwartz's distribution:

$$\dot{X}_t := \frac{d}{dt} X_t = f(t,\omega)\dot{W}.$$

On the other hand, by Theorem 2.5 we see that

$$E[X_t] = E\left[\frac{1}{2}\int_0^t \hat{f}(s,\omega)ds\right], \quad {}^\forall t.$$

We will often find it convenient to write this fact in the following form:

$$E[f(t,\cdot)\dot{W}] = \frac{1}{2}E[\hat{f}(t,\cdot)]. \tag{2.25}$$

The importance of the symmetric integral is simply explained by the following fact.

For a semi-martingale $X_t = at + bW_t + c$, (a,b,c: consts) and a smooth function $F(t,x)$, we have by the Itô formula

$$d_t F = \left\{ F_t' + aF_x' + \frac{b^2}{2} F_{xx}'' \right\} dt + bF_x' d_0 W_t.$$

With the symmetric integral this equality is expressed in the more simple form

$$d_t F = F_t' dt + F_x' \{adt + bdW_t\} = F_t' dt + F_x' dX_t,$$

since we have $\frac{\partial}{\partial W_t} F_x'(t, X_t) = F_{xx}'' b$. In other words, with the symmetric integral the differential formula in classic calculus is conserved. But the application of this property to a more general case must be done with a special care to the notion of the $B$-derivative as we see below.

### 2.3.2 Anti-Causal Function and Backward Itô Integral

This and the following Sects. 2.3.2 and 2.3.3, treat some special subjects which will be related only to a problem discussed in Chap. 7. Hence an impatient reader can skip these two subsections.

Looking back to the definition of the Itô integral $I(f),\ f \in \mathbf{S}$, we recognize that the causality condition together with the employment of the special form of Riemann sum

$$R_\Delta(f) := \sum_{t_i \in \Delta} f(t_i) \Delta_i W, \quad \Delta = \{0 \le t_1 < t_2 < \cdots < t_n \le T\}$$

is essential in endowing the martingale property to the stochastic process $I_t(f) = \int_0^t f(s, \omega) d_0 W_s$ defined by the Itô integral. A similar result might occur in a *retrograde* situation as we see below.

Let $\mathscr{G}^t := \sigma\{W_v - W_u : t \le u \le v\}$ and let $\mathscr{F}^t$ be a decreasing family of $\sigma$-fields such that

- $\mathscr{F}^t \supset \mathscr{G}^t$,
- $\mathscr{F}^t$ is independent of $\mathscr{G}_t := \sigma\{W_v - W_u : u \le v \le t\}$.

The $\sigma$-field $\mathscr{F}^t$ presents the future behaviour of the Brownian motion after time $t$, and we call the random function $f(t, \omega) \in \mathbf{H}$ *anti-causal* when it is *adapted to* the filtration $\{\mathscr{F}^t\}_t$. We will denote by $\overline{\mathbf{M}}$ the totality of anti-causal random functions, namely $\overline{\mathbf{M}} := \{f \in \mathbf{H} : f(t, \omega) \text{ is anti-causal}\}$, and by $\overline{\mathbf{M}}_2$ its subset $\overline{\mathbf{M}} \cap L^2([0, 1] \times \Omega, dt \times dP)$. We will also denote by $\overline{\mathbf{M}}_{2,c}$ the subset of $\overline{\mathbf{M}}_2$ consisting of all elements which are continuous in the mean, $\lim_{h \to 0} E[|f(t + h) - f(t)|^2] = 0$.

Given an anti-causal function $f(t, \omega)$ and a partition $\Delta = \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ of $[0, 1]$, we consider a retrograde Riemann sum

$$R^\Delta(f) := \sum_{i=0}^{n-1} f(t_{i+1}, \omega) \Delta_i W, \quad \text{where } \Delta_i W = W(t_{i+1}) - W(t_i). \tag{2.26}$$

We notice that for a causal function $f \in \mathbf{M}$ this is just the Riemann sum which leads to the integral $I_1(f)$ as $|\Delta| \to 0$. But in the present case we consider the sum for an anti-causal function. If we write

$$f^{\Delta}(t,\omega) := f(\eta^{\Delta}(t),\omega), \quad \text{where } \ \eta^{\Delta}(t) = t_{i+1} \text{ when } t \in (t_i, t_{i+1}],$$

then $f^{\Delta}(t,\omega)$ is a simple anti-causal function and the retrograde Riemann sum in (2.26) is expressed in the $I_1$-integral form;

$$R^{\Delta}(f) = I_1(f^{\Delta}).$$

The sum has the following property the verification of which is almost immediate and is omitted.

**Proposition 2.11** *When* $f \in \overline{\mathbf{M}}_2$*, we have* $E[R^{\Delta}(f)] = 0$ *and*

$$E[R^{\Delta}(f)^2] = \sum_{i=0}^{n-1} E[f^2(t_{i+1},\omega)](t_{i+1}-t_i) = E[\int_0^1 |f^{\Delta}(t,\omega)|^2 dt]. \tag{2.27}$$

Now let $\{\Delta_n\}$ be an *increasing* family of partitions such that, $\Delta_n \subset \Delta_{n+1}$ (as sets) and $\lim_{n\to\infty}|\Delta_n| = 0$. Given this and an anti-causal random function $f(t,\omega) \in \overline{\mathbf{M}}_2$ we put

$$f_n(t,\omega) := f^{\Delta_n}(t,\omega), \quad n \in \mathbf{N}.$$

Notice that for each $n$, $f_n(t,\omega) \in \overline{\mathbf{M}}_2$. We have the following statement.

**Proposition 2.12** *For an* $f(t,\omega) \in \overline{\mathbf{M}}_{2,c}$*, the sequence* $\{I_1(f_n)\}$ *converges in the mean sense.*

*Proof* By the continuity of $f(t,\omega)$ we see that $\lim_{n\to\infty} \|f - f_n\|^2 = 0$. On the other hand, by the isometry property (2.27), we find that

$$\lim_{m,n\to\infty} E[|I_1(f_m) - I_1(f_m)|^2] = \lim_{m,n\to\infty} \|f_n - f_m\|^2 = 0.$$

This completes the proof. □

**Definition 2.9** For an anti-causal function $f \in \overline{\mathbf{M}}_{2,c}$, the limit in the mean $\lim_{n\to\infty} I_1(f_n)$ of the sequence $\{I_1(f_n)\}$ in Proposition 2.12, we denote by $I_1(f)$ or by $\int_0^1 f\, d_1 W_t$, and call it the backward Itô integral.

*Remark 2.4* The same symbol $\int_0^1 f(t,\omega) d_1 W_t$ is used for the different cases, namely for the *causal* or *noncausal* functions. They are quite different from each other; for the causal function it means the sum of the Itô integral with the additional term,

$$I_1(f) = I_0(f) + \int_0^1 \hat{f}(t,\omega) dt,$$

where $\hat{f}(t,\omega) = \frac{\partial}{\partial W_t} f$, while for the anti-causal function $f(t,\omega)$ the integral $\int_0^1 f\ d_1 W_t$ is just the *backward* Itô integral. Hence when we see this notation we must be careful on the causality of the integrand.

Let $[a,b]$ be a sub-interval in $[0,1]$ and $\Delta = \{0 = t_0 < \cdots < t_n = 1\}$ be a partition of $[0,1]$. For an anti-causal function $f \in \overline{\mathbf{M}}_{2c}$ we observe that $g^{\Delta}(t,\omega) = \mathbf{1}_{[a,b]}(t) f^{\Delta}(t,\omega)$ is a simple anti-causal function and its backward Itô integral is well-defined as follows:

$$I_1(g^{\Delta}) = f(t_\ell)\{W(t_\ell) - W(a)\} + \sum_{i=\ell}^{r-1} f(t_{i+1})\Delta_i W + f(t_{r+1})\{W(b) - W(t_r)\},$$

where

$$t_\ell = \min\{t_i \geq a;\ t_i \in \Delta\}, \quad t_r = \max\{t_i \leq b;\ t_i \in \Delta\}.$$

We see by this formula that

$$E\left[|I_1(\mathbf{1}_{[a,b]} \cdot f^{\Delta})|^2\right] = E\left[\int_a^b |f^{\Delta}(t)|^2 dt\right],$$

consequently the convergence in the mean of the sequence $\{I_1(\mathbf{1}_{[a,b]} \cdot f^{\Delta})\}$ as $n \to \infty$, the limit we denote by $\int_a^b f(t,\omega) d_1 W_t$. In particular for the case $[a,b] = [t,1]$ $(0 \leq t \leq 1)$ we have

$$\begin{aligned}
&\int_t^1 f^{\Delta}(s,\omega) d_1 W_s \\
&= f(t_n)\{W(t_n) - W(t_{n-1})\} + \sum_{i=\ell+1}^{n-1} f(t_i)\Delta_i W + f(t_\ell)\{W(t_\ell) - W(t)\},
\end{aligned}$$

by which we see that the function $(t,\omega) \to \int_t^1 f^{\Delta}(s,\omega) d_1 W_s$ is adapted to the decreasing family of $\sigma$-fields $\{\mathscr{F}^t\}_t$ and that the equality

$$E[\int_t^1 f(r,\omega) d_1 W_r \mid \mathscr{F}^s] = \int_s^1 f(r,\omega) d_1 W_r, \quad P-a.s. \tag{2.28}$$

holds for any $0 \leq t \leq s$.

The integral $I_1(f)$ having been defined as the limit in the mean of the sequence of retrograde Riemann sums $\{R^1_{\Delta_n}(f)\}$, we have reached the following statement:

**Proposition 2.13** *For an anti-causal function $f \in \overline{\mathbf{M}}_{2,c}$, the function defined by the retrograde Itô integral, $\int_t^1 f\ d_1 W_r$ exhibits the martingale property of retrograde type (2.28).*

*Remark 2.5* Let $f \in \overline{\mathbf{M}}_{2,c}$ and let $Z_t(f)$ be an anti-causal process defined by

$$Z_t(f) = \exp\{\int_t^1 f(r,\omega)d_1W_r - \frac{1}{2}\int_t^1 f^2(r,\omega)dr\}.$$

Then by a similar argument we may confirm that the equality $E[Z_t(f)|\mathscr{F}^s] = Z_s(f)$ holds P-a.s. for any $t \le s$.

For the later discussion we prepare an Itô formula of backward type. Let $X_t^1,\ X_t^2$ be anti-causal Itô processes defined by

$$X_t^i = \int_t^1 f_i(s,\omega)d_1W_s + \int_t^1 g_i(s,\omega)ds, \quad i = 1, 2,$$

where $f_i,\ g_i\ (i = 1, 2)$ are anti-causal functions belonging to the class $\overline{\mathbf{M}}_{2,c}$.

Then following the same argument as in the case of causal calculus it is almost immediate to establish the next result.

**Proposition 2.14** (backward Itô formula) *For a smooth function* $F(x, y)$ *and an interval* $[a, b] \subset [0, 1]$*, the next equality holds:*

$$\begin{aligned} &F(X_b^1, X_b^2) - F(X_a^1, X_a^2) \\ &= -\int_a^b \{F_x d_1X_t^1 + F_y d_1X_t^2\} - \frac{1}{2}\int_a^b \{F_{xx} f_1^2 + F_{yy} f_2^2 + 2F_{xy} f_1 f_2\}dt. \end{aligned}$$

*Example 2.7* Let $X_t^1 = \int_t^1 f d_1W_s$ and $X_t^2 = \int_t^1 g d_1W_s$, then noting $X_1^1 = X_1^2 = 0$ we have

$$X_t^1X_t^2 = \int_t^1 \{f(s)\int_s^1 g(r)d_1W_r + g(s)\int_s^1 f(r)d_1W_r\}d_1W_s + \int_t^1 f(s)g(s)ds.$$

### 2.3.3 *The Symmetric Integral for Anti-Causal Functions*

For an anti-causal function $X(t,\omega)$ its symmetric integral (of backward type) can be defined similarly to the case of causal functions. Given a partition $\Delta = \{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = 1\}$ we consider for a fixed $\theta \in [0, 1]$ the Riemann sum as follows:

$$R_\Delta^\theta(f) := \sum_{k=1}^n X(t_k + \theta\Delta_k)\Delta_k W, \tag{2.29}$$

where, $\Delta_k = t_{k+1} - t_k,\ \ \Delta_k W = W(t_{k+1}) - W(t_k)$.

We know that the sequence $R_\Delta^1(f)$ converges as $|\Delta| \to 0$, and to assure the convergence for the case $\theta < 1$ we need some assumption on the regularity of the integrand $X_t$, namely a kind of B-differentiability. But for the simplicity of discussion we suppose that the anti-causal function $X_t$ is given in the formula of the backward Itô integral,

$$X_t = \int_t^1 f(s,\omega) d_1 W_s + \int_t^1 g(s,\omega) ds \qquad f,\ g \in \overline{\mathbf{M}}_{2,c}.$$

We may call such a process the *backward* Itô process.

**Proposition 2.15** *Let* $\{\Delta_n\}$ *be an increasing family of partitions of* $[0,1]$*. Then for each fixed* $\theta \in [0,1]$ *the sequence of retrograde Riemann sums* $\{R^{\theta}_{\Delta_n}(f)\}$ *converges in the mean as* $|\Delta_n| \to 0$*, and the limit which we denote by* $\int_0^1 X_s d_\theta W_s$ *is expressed in the following form:*

$$\int_0^1 X_s \, d_\theta W_s = \int_0^1 X_s \, d_1 W_s + \theta \int_0^1 f(s,\omega) ds. \tag{2.30}$$

*Proof* Let us write $R_n(\theta) = R^{\theta}_{\Delta_n}(f)$ and $t_k(\theta) = t_k + \theta \Delta_k$, then we have

$$\begin{aligned} R_n(\theta) - R_n(1) &= \sum_{k=1}^{n} \{X(t_k + \theta\Delta_k) - X(t_{k+1})\}\Delta_k W \\ &= \sum_{k=0}^{n-1} \int_{t_k(\theta)}^{t_{k+1}} f(s,\omega) d_1 W_s \cdot \Delta_k W. \end{aligned}$$

By the formula in Example 2.7 we find

$$\begin{aligned} &\int_{t_k(\theta)}^{t_{k+1}} f(s,\omega) d_1 W_s \cdot \Delta_k W \\ &= \int_{t_k(\theta)}^{t_{k+1}} f(s,\omega)\{W(t_{k+1}) - W(s)\} d_1 W_s + \int_{t_k}^{t_{k+1}} X_s d_1 W_s + \int_{t_k(\theta)}^{t_{k+1}} f(s.\omega) ds\} \\ &= T_1(k) + T_2(k), \end{aligned}$$

where

$$\begin{aligned} T_1(k) &= \int_{t_k(\theta)}^{t_{k+1}} f(s,\omega)\{W(t_{k+1}) - W(s)\} d_1 W_s + \int_{t_k}^{t_{k+1}} X_s d_1 W_s, \\ T_2(k) &= \int_{t_k(\theta)}^{t_{k+1}} f(s.\omega) ds. \end{aligned}$$

It is routine to verify that

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} T_1(k) = 0 \ \text{ in } \ L^2(\Omega, dP)$$

and that

$$\lim_{n\to\infty}\sum_{k-1}^{n} T_2(k) = \theta \int_0^1 f(s,\omega)ds.$$

Since $\lim_{n\to\infty} R_n(1) = \int_0^1 X_s\, d_1 W_s$, this completes the proof. $\square$

## 2.4 SDE

A stochastic functional equation for an unknown process $X_t$ as follows is called a stochastic integral equation,

$$X_t = x_0 + \int_0^t b(s, X_s)ds + \int_0^t a(s, X_s)d_0 W_s, \qquad 0 \le t \le T, \tag{2.31}$$

where $x_0 \in \mathbf{R}$ and $a(t,x), b(t,x)$ are real functions measurable in $(t,x)$. In the condition that the unknown process $X(t,\omega)$ is limited to be causal (with respect to the filtration $\{\mathscr{F}_t^W\}$) the equation becomes meaningful in the framework of the Itô integral. It is customary to represent this Eq. (2.31) by the following symbolic form which is called the stochastic differential equation (or SDE for short) of Itô type:

$$dX_t = b(t, X_t)dt + a(t, X_t)d_0 W_t, \quad X_0 = x_0. \tag{2.32}$$

The discussion on SDEs based on Itô calculus is not our principal subject in this book. So we do not give here a detailed review about it, but we intend to give only some elementary results for the reference in later chapters.

### *2.4.1 Strong Solution*

**Definition 2.10** A continuous stochastic process $X_t(\omega)$, $(t \ge 0)$, defined on the same probability space $(\Omega, \mathscr{F}, P)$ as Brownian motion $W_t(\omega)$ and adapted to the filtration $\{\mathscr{F}_t^W, t \ge 0\}$, is called the *strong solution* of the SDE (2.32) provided that the couple $(W, X)$ satisfies the Eq. (2.31) with probability one for all $t \in [0, T]$.

As for the fundamental properties of the strong solution we have the following statement.

**Theorem 2.6** *Let $a(t,x)$, $b(t,x)$ be real and smooth functions with bounded derivatives in $x$, i.e. $|\partial_x b(t,x)|$, $|\partial_x a(t,x)| <^{\exists} L_0 \ {}^{\forall}(t,x)$. We suppose that the initial data $X_0(\omega)$ is independent of the Brownian motion and $E[X_0^2] < \infty$. Then the Cauchy problem (2.32) for SDE, or equivalently the SIE (2.31) has the unique strong solution $X_t$.*

*Proof* Put $L = \max\{L_0, \sup_{t\in[0,T]} |b(t,0)|, \sup_{t\in[0,1]} |a(t,0)|\}$, then by the assumption on the coefficients $a(t,x), b(t,x)$, we notice that they satisfy the following conditions:

$$|b(t,x) - b(t,y)| < L|x-y|,\ |a(t,x) - a(t,y)| < L|x-y| \text{ and}$$
$$|a(t,x)|, |b(t,x)| < L\sqrt{1+x^2} \text{ for any } x. \tag{2.33}$$

We are going to prove the statement by the standard method of Picard. So we construct a sequence of random functions $\{X_n(t)\}$ in the following way:

$$X_0(t,\omega) = X_0(\omega)$$
$$X_{n+1}(t,\omega) = X_0(\omega) + \int_0^t b(s, X_n(s))ds + \int_0^t a(s, X_n(s))d_0W_s, \quad \text{for } n \geq 1. \tag{2.34}$$

First of all we notice that the assumption $E[X_0(t,\omega)^2] < \infty$ implies by the second condition in (2.33) that $X_1(t,\omega) \in \mathbf{M}_2$, hence by induction we notice that every $X_n$ is well-defined as an element of $\mathbf{M}_2$.

From definition (2.34) we have

$$X_n(t) - X_{n-1}(t) = \int_0^t \alpha_{n-1}(s,\omega)ds + \int_0^t \beta_{n-1}(s,\omega)d_0W_s, n \geq 1, \tag{2.35}$$

where

$$\alpha_n(s,\omega) = b(s, X_n(s)) - b(s, X_{n-1}(s)), \quad \beta_n(s,\omega) = a(s, X_n(s)) - a(s, X_{n-1}(s)).$$

As for these we notice that

$$|\alpha_n(s,\omega)|,\ |\beta_n(s,\omega)| \leq L|X_n(s) - X_{n-1}(s)| \quad \text{for all } (s,\omega). \tag{2.36}$$

We put $d_n(t) = E[|X_n(t) - X_{n-1}(t)|^2]$ and we claim that

$$d_n(t) \leq C_1 \frac{(C_2 t)^n}{n!} \quad \text{for some constants } C_1, C_2. \tag{2.37}$$

For $n = 1$ we have

$$X_1(t) - X_0(t) = \int_0^t b(s, X_0(\omega))ds + \int_0^t a(s, X_0(\omega))d_0W_s,$$

from which we easily find that

$$d_1(t) \leq C_1C_2t \text{ where } C_1 = 1 + E[|X_0|^2], C_2 = 2(T+1)L^2. \tag{2.38}$$

For the case $n \geq 2$, with the help of the condition (2.36) and Schwarz inequality, we get from (2.34) the following inequality:

$$\begin{aligned}
d_n(t) &= E[|X_n(t) - X_{n-1}(t)|^2] \\
&\leq 2E\left[\left(\int_0^t \alpha_{n-1}(s,\omega)ds\right)^2 + (\beta_{n-1}(s,\omega)d_0W_s)^2\right] \\
&\leq 2L^2\left[T\int_0^t E[|X_{n-1}(s) - X_{n-1}(s)|^2]ds + \int_0^t E[|X_{n-1}(s) - X_{n-2}(s)|^2]ds\right] \\
&= 2(T+1)L^2\int_0^t d_{n-1}(s)ds.
\end{aligned}$$

Namely we find

$$d_n(t) \leq C_2\int_0^t d_{n-1}(s)ds.$$

By induction from this integral inequality, together with the estimate (2.38), we get the desired estimate (2.37).

Next we show the uniform convergence of the sequence $\{X_n(t)\}$. Again from the equality (2.34), we have

$$\max_{t\in[0,T]}|X_{n+1}(t) - X_n(t)| \leq \max_{t\in[0,T]}\left\{|\int_0^t \alpha_n(s)ds| + |\int_0^t \beta_n(s)d_0W_s|\right\}.$$

Thus for any fixed $M > 0$, we have the following inequality:

$$\begin{aligned}
&P\{\max_{t\in[0,T]}|X_{n+1}(t) - X_n(t)| > 2M\} \\
&\leq P\left\{\max_{t\in[0,T]}\left|\int_0^t \alpha_n(s)ds\right| > M\right\} + P\left\{\max_{t\in[0,T]}\left|\int_0^t \beta_n(s)d_0W_s\right| > M\right\}. \qquad (2.39)
\end{aligned}$$

As for the first term on the right hand side, we have

$$\begin{aligned}
&P\left\{\max_{t\in[0,T]}\left|\int_0^t \alpha_n(s)ds\right| > M\right\} \leq P\left\{T\int_0^T \alpha_n^2(s)ds > M^2\right\} \\
&\leq \frac{L^2T}{M^2}E\int_0^T d_{n-1}(s)ds \leq C_3\frac{(C_1T)^n}{n!M^2},
\end{aligned}$$

where $C_3 = \frac{C_2L^2T}{C_1}$.

As for the second term, by applying Doob's submartingale inequality (see Theorem 10.2 in Appendices) we get the following estimate:

$$P\left\{\max_{t\in[0,T]}\left|\int_0^t \beta_n(s)d_0W_s\right| \geq M\right\} \leq \frac{1}{M^2}E\left[\left(\int_0^T \beta_n(s)d_0W_s\right)^2\right]$$
$$\leq \frac{L^2}{M^2}\int_0^T d_{n-1}(s)ds \leq C_4\frac{(C_1T)^n}{n!M^2},$$

where $C_4 = \max\{C_3, (C_2L^2/C_1)\}$.

Combining these two estimates with the inequality (2.39) and putting $M = \dfrac{1}{\sqrt[4]{n!}}$, we find

$$P\left[\max_{t\in[0,T]}|X_{n+1}(t) - X_n(t)| > \frac{2}{\sqrt[4]{n!}}\right] \leq 2C_4\frac{(C_1T)^n}{\sqrt{n!}}.$$

The series $\sum_n \frac{(C_1T)^n}{\sqrt{n!}}$ being convergent, by virtue of Borel- Cantelli's first lemma we get from the above inequality the following result:

$$P\left[\max_{t\in[0,T]}|X_{n+1}(t) - X_n(t)| \leq \frac{2}{\sqrt[4]{n!}} \text{ for large enough } n\right] = 1.$$

This implies that the sequence $\{X_n(t,\omega)\}$ converges to a limit $X_\infty(t)$ almost surely and uniformly in $t \in [0, T]$. Now letting $n \to \infty$ on both sides of the equation (2.34) we confirm that the limit $X_\infty(t)$ solves the SDE.

What is left is the verification of the uniqueness of strong solution. So let $Y(t,\omega)$ be another strong solution of the SDE, for which we have the following equality:

$$X(t) - Y(t) = \int_0^t \{b(s, X(s)) - b(s, Y(s))\}ds + \int_0^t \{a(s, X(s)) - a(s, Y(s))\}d_0W_s.$$

Put $d(t) = E[|X(t) - Y(t)|^2]$ then, following the same argument as we have done, we find that

$$d(t) \leq 2L^2(1+T)\int_0^t d(s)ds.$$

Since $d(0) = 0$ the application of Gronwall's inequality (see the subject in Chap. 10) shows us that $d(t) = 0$ for any $t$, hence

$$P\{X(t) = Y(t)\} = 1 \text{ for any } t.$$

This completes the proof. □

*Remark 2.6* The solutions $X(t), Y(t)$ being continuous, we see by separability of those processes that $P\{X(t) = Y(t) \ ^\forall t \in [0, T]\} = 1$.

### 2.4.2 Law of the Solution of SDE

We have shown a statement on the existence and uniqueness of the strong solution of the Cauchy problem for the SDE (2.32). It would be intuitively clear that this solution $X_t$ is a Markov process, for the following two reasons: (1) Because of the uniqueness as the solution of the SDE, the value of the process after "$t$" $\{X_u,\ u \geq t\}$ depends only on the final data "$X_t$" and the increments $\{W_u - W_t,\ u \geq t\}$ of the driving force $W_{\cdot}$; besides (2) the increments $\{W_u - W_t,\ u \geq t\}$ are independent of the past history $\mathscr{F}^W_{t-}$.

Hence we are interested in the transition probability of the solution $X$ of the SDE,

$$P(s, x, t, dy) = P\{X_t \in dy | X_s = x\}, \quad s \leq t,\ x \in \mathbf{R}^1.$$

Suppose for the simplicity of discussion that the transition kernel has the density $P(s, x, t, dy) = p(s, x, t, y)dt$. Now fix a smooth function $f(x) \in C_b^2$ with finite support and consider the expectation

$$E[f(X_t)|X_s = x] = \int_{\mathbf{R}^1} f(y)p(s, x, t, y)dy.$$

By the Itô formula we have the equality

$$f(X_t) = f(x) + \int_s^t f'(X_r)\{b(r, X_r)ds + a(r, X_r)d_0 W_r\} + \frac{1}{2}\int_0^t f''(X_r)b^2(r, X_r)ds,$$

from which we see that,

$$E[f(X_t)|X_s = x] = f(x) + E\left[\int_s^t \{f'(X_r)b(r, X_r) + \frac{1}{2}f''(X_r)b^2(r, X_r)\}ds\right].$$

By changing the order of integrations we find the following equality:

$$\begin{aligned}
&\int_{\mathbf{R}^1} f(y)p(s, x, t, y)dy \\
&= f(x) + \int_s^t dr \int_{\mathbf{R}^1} p(s, x, t, y)dy\{f'(y)b(r, y) + \frac{1}{2}f''(y)b^2(r, y)\}.
\end{aligned}$$

Thus by taking into account the fact that $f(x)$ is of compact support and by applying the integration by parts formula to this, we get the following:

$$\begin{aligned}
\int_{\mathbf{R}^1} f(y)dy \int_s^t dr \Big[&\partial_r p(s, x, r, y) + \partial_y\{b(r, y)p(s, x, r, y)\} \\
&- \frac{1}{2}\partial_y^2\{b^2(r, y)p(s, x, r, y)\}\Big] = 0.
\end{aligned}$$

The test function $f(x)$ being arbitrary, by virtue of Weyl's lemma (see for example [15]) we get from this equality the equation for $u(t,x) := p(0,x_0,t,x)$ as follows:

$$\begin{aligned}&\frac{\partial}{\partial t}u(t,x) = -\frac{\partial}{\partial x}\{b(t,x)u(t,x)\} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\{b^2(t,x)u(t,x)\},\\ &u(0,x) = \delta_{x_0}(x).\end{aligned} \tag{2.40}$$

This is the so-called *Kolmogorov forward equation.*

The *backward equation* can be obtained by taking $u(t,x) = E[f(X_T)|X_t = x]$ and applying a similar argument based on the Itô formula, which would read as follows:

$$\begin{aligned}&\frac{\partial}{\partial t}u(t,x) + b(t,x)\frac{\partial}{\partial x}u(t,x) + \frac{b^2(t,x)}{2}\frac{\partial^2}{\partial x^2}u(t,x) = 0,\ t \le T,\\ &u(T,x) = f(x).\end{aligned} \tag{2.41}$$

In terms of the transition probability density $p(t,x,T,y)$ the function $u(t,x)$ is written as $u(t,x) = \int_{\mathbf{R}^1} f(y)p(t,x,T,y)dy$, hence again by applying Weyl's lemma we see that from equations in (2.41) the following equations hold:

$$\begin{aligned}&\{\frac{\partial}{\partial t} + b(t,x)\frac{\partial}{\partial x} + \frac{b^2(t,x)}{2}\frac{\partial^2}{\partial x^2}\}p(t,x,T,y) = 0,\ t \le T,\\ &p(T,x,T,y) = \delta_y(x).\end{aligned}$$

### 2.4.3 Martingale $Z_t$ and Girsanov's Theorem

For a nice real function $f(t,\omega)$ belonging to the class **M**, we have introduced the causal function $Z_t$ by the following form:

$$Z_t = \exp\left\{\int_0^t f(s,\omega)d_0W_s - \frac{1}{2}\int_0^t f^2(s,\omega)ds\right\}.$$

We have seen that this positive function is the unique strong solution of the Itô SDE

$$dZ_t = f(t,\omega)Z_t d_0 W_t, \qquad Z_0 = 1.$$

By this fact we notice that the $Z_t$ is an $\mathscr{F}_t^W$-martingale with $E[Z_t] = E[Z_0] = 1$. Moreover since $Z_t > 0$ we see that $dQ = Z_t dP$ becomes another probability on the same measurable space $(\Omega, \mathscr{F})$. We denote the expectation with respect to this new probability by $E^Q[\cdot]$. It is interesting to ask how the Brownian motion $W_t$ looks like under new probability $dQ$. For this aim we consider a stochastic process $Y_t$ as follows: $dY_t = dW_t - f(t,\omega)dt$.

Then by Itô formula we obtain the equality,

$$
\begin{aligned}
d(Y_t Z_t) &= Z_t\{d_0 W_t - f(t,\omega)dt\} + Y_t f(t,\omega) Z_t d_0 W_t + Z_t f(t,\omega)dt \\
&= Z_t\{1 + Y_t f(t,\omega)\} d_0 W_t,
\end{aligned}
$$

which shows that $Y_t Z_t$ is an $\mathscr{F}_t^W$-martingale under the original measure $dP$. Hence for an arbitrary event $A \in \mathscr{F}_s^W$, we have,

$$
E^Q[\mathbf{1}_A Y_t] = E[\mathbf{1}_A Y_t Z_t] = E[\mathbf{1}_A E[Y_t Z_t | \mathscr{F}_s^W]] = E[\mathbf{1}_A Y_s Z_s] = E^Q[\mathbf{1}_A Y_s],
$$

in other words, $E^Q[Y_t|\mathscr{F}_s^W] = Y_s$ Q-a.s.. This means that $Y_t$ is an $\mathscr{F}_t^W$-martingale under the measure $dQ$.

On the other hand, we easily see that the quadratic variation of $Y_t$ is $d[Y]_t = dt$. Consequently by Lemma 9.3 we get the following result:

**Proposition 2.16** (Girsanov's Theorem) *Under the measure $dQ = Z_t dP$ the process $Y_t$; $dY_t = dW_t - f(t,\omega)dt$ becomes a Brownian motion.*

http://www.springer.com/978-4-431-56574-1

Noncausal Stochastic Calculus
Ogawa, S.
2017, XII, 210 p. 1 illus., Hardcover
ISBN: 978-4-431-56574-1