Chapter 2. Brownian motion calculus

2.1 Introduction

Let $B(t)$ be a Brownian motion, together with a filtration $\mathscr{F}_t, t \geq 0$. Our goal to define stochastic integral

$$\int_0^T X(t)dB(t),$$

the integrand $X(t)$ can also be a stochastic process. The integral should be well-defined for at least all non-random continuous functions on $[0, T]$. When the integrand is random, we will assume that it is an adapted stochastic process.

Riemann-Stieltjes integral

Stieltjes integral of a function f with respect to a function g over interval $(a,b]$ is defined as

$$\int_a^b f dg = \int_a^b f(t)dg(t) = \lim_{\delta} \sum_{i=1}^{n} f(\xi_i^n)\big[g(t_i^n) - g(t_{i-1}^n)\big],$$

where $\{t_i^n\}$ represent partitions of the interval,

$$a = t_0^n < t_1^n < \ldots < t_n^n = b, \ \ \delta = \max_{1\le i\le n}(t_i^n - t_{i-1}^n),$$

and $\xi_i^n \in [t_{i-1}^n, t_i^n]$,

the limit is taken over all partitions and all choice of ξ_i^n's with $\delta \to 0$.

When g is a function of finite variation, then any continuous function function f is stieltjes-integrable with respect to g.

Theorem

Let $\delta = \max_{1\le i\le n}(t_i^n - t_{i-1}^n)$ *denote the length of the largest interval in partition of* $[a,b]$. *If*

$$\lim_{\delta} \sum_{i=1}^{n} f(\xi_i^n)\left[g(t_i^n) - g(t_{i-1}^n)\right]$$

exists for any continuous function f, *then* g *must be of finite variation on* $[a,b]$.

This shows that if g (for example, the Brownian motion B) has infinite variation then the limit of the approximating sums may not exist even when the integrand function f is continuous. Therefore integrals with respect to functions of infinite variation (stochastic integrals) must be defined in another way.

2.2 Definition of Itô integral

The integral integral $\int_0^T X(t)dB(t)$ should have the properties:

- If $X(t) = 1$, then $\int_0^T X(t)dB(t) = B(T) - B(0)$;
- If $X(t) = c$ in $(a, b] \subset [0, T]$ and zero otherwise, then
$$\int_0^T X(t)dB(t) = c(B(b) - B(a));$$
- For real α and β,
$$\int_0^T \big(\alpha X(t) + \beta Y(t)\big)dB(t)$$
$$= \alpha \int_0^T X(t)dB(t) + \beta \int_0^T Y(t)dB(t).$$

2.2.1 Itô's Integral for simple integrands

Deterministic simple process

Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, T]$; i.e.,

$$0 = t_0 < t_1 < \ldots < t_n = T.$$

Assume that $X(t)$ is a non-random constant in t on each subinterval $(t_i, t_{i+1}]$, i.e.,

$$X(t) = \begin{cases} c_0 & \text{if } t = 0, \\ c_i & \text{if } t_i < t \le t_{i+1}, i = 0, \ldots, n-1. \end{cases}, \qquad (2.2.1)$$

or in one formula,

$$X(t) = c_0 I_0(t) + \sum_{i=0}^{n-1} c_i I_{(t_i, t_{i+1}]}(t).$$

Such a process is a deterministic simple process. The Itô integral $\int_0^T X(t)dB(t)$ is defined as a sum

$$\int_0^T X(t)dB(t) = \sum_{i=0}^{n-1} c_i \big(B(t_{i+1}) - B(t_i)\big).$$

It is easy to see that the integral is a Gaussian random variable with mean zero and variance

$$\mathrm{Var}\Big(\int XdB\Big) = \sum_{i=0}^{n-1} c_i^2 \mathrm{Var}\big(B(t_{i+1}) - B(t_i)\big)$$
$$= \sum_{i=0}^{n-1} c_i^2 (t_{i+1} - t_i) = \int_0^T X^2(t)dt.$$

Simple process

Since we would like to integrate random process, it is important to allow constants c_i in (2.2.1) to be random. If c_i's are replaced by random variables ξ_i's, then in order to carry out calculations, and have convenient properties of the integral, the random variable ξ_i are allowed to depend on the values of $B(t)$ for $t \le t_i$, that is, they are allowed to be $\mathscr{F}_{t_i}$-measurable, and independent of the future increments of the Brownian motion.

In finance, $X(t)$ is regraded as an asset at time t, it depends on the information of the past. So, $X(t)$ is assumed to be adapted at least. So, ξ_i is $\mathscr{F}_{t_i}$- measurable.

A simple process is

$$X(t) = \xi_0 I_0(t) + \sum_{i=0}^{n-1} \xi_i I_{(t_i, t_{i+1}]}(t), \text{ with } \xi_i \text{ is } \mathscr{F}_{t_i} - \text{measurable.}$$

A simple process is

$$X(t) = \xi_0 I_0(t) + \sum_{i=0}^{n-1} \xi_i I_{(t_i, t_{i+1}]}(t), \text{ with } \xi_i \text{ is } \mathscr{F}_{t_i} - \text{measurable.}$$

For a simple process, the Itô integral $\int_0^T X dB$ is defined as a sum

$$\int_0^T X(t) dB(t) = \sum_{i=0}^{n-1} \xi_i \big(B(t_{i+1}) - B(t_i) \big). \tag{2.2.2}$$

In general, for $0 \leq t \leq T$, if $t_k < t \leq t_{k+1}$, then

$$\int_0^t X(u)dB(u) = \sum_{j=0}^{k-1} \xi_j \big(B(t_{j+1}) - B(t_j)\big) + \xi_k \big(B(t) - B(t_k)\big)$$

$$= \sum_j \xi_j \big(B(t \wedge t_{j+1}) - B(t \wedge t_j)\big).$$

Properties of Itô integral of simple processes

(i) Linearity. If $X(t)$ and $Y(t)$ are simple processes and α and β are constant, then

$$\int_0^T \big(\alpha X(t) + \beta Y(t)\big) dB(t)$$

$$= \alpha \int_0^T X(t) dB(t) + \beta \int_0^T Y(t) dB(t).$$

Properties of Itô integral of simple processes

(i) Linearity. If $X(t)$ and $Y(t)$ are simple processes and α and β are constant, then

$$\begin{aligned}&\int_0^T \big(\alpha X(t)+\beta Y(t)\big)dB(t)\\ =&\alpha\int_0^T X(t)dB(t)+\beta\int_0^T Y(t)dB(t).\end{aligned}$$

(ii) For all $[a,b]\subset[0,T]$,

$$\int_0^T I_{[a,b]}(t)dB(t)=B(b)-B(a).$$

The next properties hold if ξ_i's are square integrable, $\mathsf{E}\xi_i^2<\infty$.

(iii) Zero mean property.

$$\mathsf{E}\int_0^T X(t)dB(t) = 0.$$

(iii) Zero mean property.

$$\mathsf{E}\int_0^T X(t)dB(t) = 0.$$

(iv) Isometry property.

$$\mathsf{E}\left(\int_0^T X(t)dB(t)\right)^2 = \int_0^T \mathsf{E}[X^2(t)]dt.$$

(v) Martingale property. Let $I(t) = \int_0^t X(u)dB(u)$.

Then $I(t)$ is a continuous martingale.

(v) Martingale property. Let $I(t) = \int_0^t X(u)dB(u)$. Then $I(t)$ is a continuous martingale.

(vi) The quadratic variation accumulated up to time t by the Itô integral is

$$[I, I](t) = \int_0^t X^2(u)du.$$

(v) Martingale property. Let $I(t)=\int_0^t X(u)dB(u)$. Then $I(t)$ is a continuous martingale.

(vi) The quadratic variation accumulated up to time t by the Itô integral is

$$[I,I](t)=\int_0^t X^2(u)du.$$

(vii) $I^2(t)-\int_0^t X^2(u)du$ is a martingale, or equivalently, for all $s\le t$,

$$\mathsf{E}\left[\left(\int_s^t X(u)dB(u)\right)^2\Big|\mathscr{F}_s\right]=\int_s^t \mathsf{E}\left[X^2(u)\big|\mathscr{F}_s\right]du.$$

Proof. Property (i) and (ii) are easy to be verified directly from the definition. Property (iii) follows from Property (v).

Property (iv). Write $D_i = B(t_{i+1}) - B(t_i))$. Then

$$(I(T))^2 = (\sum_{i=0}^{n-1} \xi_i D_i)^2 = \sum_{i=0}^{n-1} \xi_i^2 D_i^2 + 2 \sum_{j<i} \xi_i \xi_j D_i D_j.$$

Property (iv). Write $D_i = B(t_{i+1}) - B(t_i))$. Then

$$(I(T))^2 = (\sum_{i=0}^{n-1} \xi_i D_i)^2 = \sum_{i=0}^{n-1} \xi_i^2 D_i^2 + 2\sum_{j<i} \xi_i \xi_j D_i D_j.$$

Notice that ξ_i is $\mathscr{F}_{t_i}$-measurable and $D_i = B(t_{i+1}) - B(t_i)$ is independent of $\mathscr{F}_{t_i}$. So

$$\mathsf{E}[\xi_i^2 D_i^2 | \mathscr{F}_{t_i}] = \xi_i^2 \mathsf{E}[D_i^2] = \xi_i^2 (t_{i+1} - t_i).$$

It follows that

$$\mathsf{E}[[\xi_i^2 D_i^2] = \mathsf{E}\left[\mathsf{E}[\xi_i^2 D_i^2 | \mathscr{F}_{t_i}]\right] = \mathsf{E}[\xi_i^2](t_{i+1} - t_i).$$

Also, for $j < i$, ξ_j, D_j are $\mathscr{F}_{t_i}$-measurable. So

$$\mathsf{E}[\xi_i \xi_j D_i D_j | \mathscr{F}_{t_i}] = \xi_j D_j \xi_i \mathsf{E}[D_i] = 0.$$

It follows that

$$\mathsf{E}[\xi_i \xi_j D_i D_j] = 0.$$

Hence

$$\mathsf{E}(I(T))^2 = \sum_{i=0}^{n-1} \mathsf{E}[\xi_i^2](t_{i+1} - t_i) = \int_0^T \mathsf{E}(X^2(t))dt.$$

Property (v). For $0 \leq s \leq t \leq T$, we want to show that $\mathsf{E}[I(t)|\mathscr{F}_s] = I(s)$. It suffices to show that $\mathsf{E}[I(T)|\mathscr{F}_s] = I(s)$, which implies

$$\mathsf{E}[I(t)|\mathscr{F}_s] = \mathsf{E}\left[\mathsf{E}[I(T)|\mathscr{F}_t]\Big|\mathscr{F}_s\right] = \mathsf{E}[I(T)|\mathscr{F}_s] = I(s).$$

Assume $s \in (t_k, t_{k+1}]$, write $I(T)$ as

$$I(T) = \sum_{i=0}^{k-1} \xi_i D_i + \xi_k[B(t_{k+1}) - B(t_k)] + \sum_{i=k+1}^{n-1} \xi_i D_i.$$

For $i \le k-1$, $t_{i+1} \le s$, ξ_i and D_i are $\mathscr{F}_{t_{i+1}}$-measurable, and so $\mathscr{F}_s$-measurable. It follows that

$$\mathsf{E}\left[\sum_{i=0}^{k-1} \xi_i D_i \middle| \mathscr{F}_s\right] = \sum_{i=0}^{k-1} \xi_i D_i.$$

For $i \le k-1$, $t_{i+1} \le s$, ξ_i and D_i are $\mathscr{F}_{t_{i+1}}$-measurable, and so $\mathscr{F}_s$-measurable. It follows that

$$\mathsf{E}\left[\sum_{i=0}^{k-1}\xi_i D_i \Big| \mathscr{F}_s\right] = \sum_{i=0}^{k-1}\xi_i D_i.$$

For $i \ge k+1$, $\mathscr{F}_{t_i} \supset \mathscr{F}_s$ and $D_i = B(t_{i+1}) - B(t_i)$ is independent of $\mathscr{F}_{t_i}$. So

$$\mathsf{E}\left[\xi_i D_i \Big| \mathscr{F}_s\right] = \mathsf{E}\left[\mathsf{E}\left[\xi_i D_i | \mathscr{F}_{t_i}\right] \Big| \mathscr{F}_s\right] = \mathsf{E}\left[\xi_i \mathsf{E}[D_i] \Big| \mathscr{F}_s\right] = 0.$$

Finally,

$$\begin{aligned}\mathsf{E}\left[\xi_k[B(t_{k+1})-B(t_k)]|\mathscr{F}_s\right] =&\xi_k[\mathsf{E}\left[B(t_{k+1})|\mathscr{F}_s\right]-B(t_k)]\\ =&\xi_k[B(s)-B(t_k)].\end{aligned}$$

Finally,

$$\begin{aligned} \mathsf{E}\left[\xi_k[B(t_{k+1}) - B(t_k)]|\mathscr{F}_s\right] =& \xi_k[\mathsf{E}\left[B(t_{k+1})|\mathscr{F}_s\right] - B(t_k)] \\ =& \xi_k[B(s) - B(t_k)]. \end{aligned}$$

It follows that

$$\mathsf{E}[I(T)|\mathscr{F}_s] = \sum_{i=0}^{k-1} \xi_i D_i + \xi_k[B(s) - B(t_k)] = I(s).$$

Property (vi). Assume $t_k \le t \le t_{k+1}$. Notice,

$$I(s) = \sum_{j=0}^{i-1} \xi_j \big(B(t_{j+1}) - B(t_j)\big) + \xi_i \big(B(s) - B(t_i)\big), \quad \text{for } t_i < s \le t_{i+1}$$

Property (vi). Assume $t_k \le t \le t_{k+1}$. Notice,

$$I(s) = \sum_{j=0}^{i-1} \xi_j \big(B(t_{j+1}) - B(t_j)\big) + \xi_i \big(B(s) - B(t_i)\big), \quad \text{for } t_i < s \le t_{i+1}$$

It follows that

$$[I, I]([t_i, t_{i+1}]) = \xi_i^2 \cdot [B, B]([t_i, t_{i+1}]) = \xi_i^2 (t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} X^2(u) du.$$

Similarly,

$$[I, I]([t_k, t]) = \xi_k^2 \cdot [B, B]([t_k, t]) = \xi_k^2 (t - t_k) = \int_{t_k}^{t} X^2(u) du.$$

Adding up all these pieces, we obtain

$$[I, I](t) = [I, I]([0, t]) = \int_0^t X^2(u) du.$$

Property (vii). Let $M(t)=I^2(t)-[I,I](t)$. It suffices to show that $\mathsf{E}[M(T)|\mathscr{F}_s]=M(s)$. Assume $s\in(t_k,t_{k+1}]$. Then

$$\int_s^T X(u)dB(u)=\xi_k\big[B(t_{k+1})-B(s)\big]+\sum_{j=k+1}^{n-1}\xi_j\big[B(t_{j+1})-B(t_j)\big].$$

Following the same lines in the proof of Property 4, we have

$$\mathsf{E}\left[\left(\int_s^T X(u)dB(u)\right)^2\Big|\mathscr{F}_s\right]$$

$$=\mathsf{E}[\xi_k^2|\mathscr{F}_s](t_{k+1}-s)+\sum_{j=k+1}^{n-1}\mathsf{E}[\xi_j^2|\mathscr{F}_s](t_{j+1}-t_j)$$

$$=\int_s^T \mathsf{E}\big[X^2(u)|\mathscr{F}_s\big]du.$$

Hence

$$
\begin{aligned}
&\mathsf{E}\left[I^2(T)\middle|\mathscr{F}_s\right]\\
=&I^2(s)+2I(s)\mathsf{E}\left[I(T)-I(s)\middle|\mathscr{F}_s\right]+\mathsf{E}\left[\left(\int_s^T X(u)dB(u)\right)^2\middle|\mathscr{F}_s\right]\\
=&I^2(s)+\int_s^T \mathsf{E}\left[X^2(u)|\mathscr{F}_s\right]du.
\end{aligned}
$$

On the other hand, it is obvious that

$$\mathsf{E}\left[[I,I](T)\middle|\mathscr{F}_s\right] = [I,I](s) + \int_s^T \mathsf{E}\left[X^2(u)|\mathscr{F}_s\right]du.$$

It follows that

$$\mathsf{E}\left[I^2(T) - [I,I](T)\middle|\mathscr{F}_s\right] = I^2(s) - [I,I](s),$$

that is $\mathsf{E}[M(T)|\mathscr{F}_s] = M(s)$.

2.2.2 Itô's Integral for general integrands

Let $X^n(t)$ be a sequence of simple processes convergent in probability to the process $X(t)$. Then, under some conditions, the sequence of their integrals $\int_0^T X^n(t)dB(t)$ also convergence in probability. That limit is taken to be the integral $\int_0^T X(t)dB(t)$.

Example

Find $\int_0^T B(t)dB(t)$.

Example

Find $\int_0^T B(t)dB(t)$.

Solution. Let $0 = t_0^n < t_1^n < \ldots < t_n^n = T$ be a partition of $[0, T]$, and let

$$X^n(t) = \sum_{i=0}^{n-1} B(t_i^n) I_{(t_i^n, t_{i+1}^n]}(t).$$

Then, for each n, X^n is a simple process.

Take the sequence of partitions such that $\delta_n = \max_i (t_{i+1}^n - t_i^n) \to 0$. Then $X^n(t) \to B(t)$ almost surely, be the continuity of the Brownian paths.

Take the sequence of partitions such that $\delta_n = \max_i (t_{i+1}^n - t_i^n) \to 0$. Then $X^n(t) \to B(t)$ almost surely, be the continuity of the Brownian paths.

We will show that $\int_0^T X^n(t) dB(t)$ will converges in probability.

$$\int_0^T X^n(t)dB(t) = \sum_{i=0}^{n-1} B(t_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big).$$

$$\int_0^T X^n(t)dB(t) = \sum_{i=0}^{n-1} B(t_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big).$$

Observer

$$\begin{aligned} & B(t_i^n)\big(B(t_{i+1}^n) - B(t_i^n)\big) \\ =& \frac{1}{2}\Big(B^2(t_{i+1}^n) - B^2(t_i^n) - \big(B(t_{i+1}^n) - B(t_i^n)\big)^2\Big), \end{aligned}$$

and then

$$\int_0^T X^n(t)dB(t)$$

$$=\sum_{i=0}^{n-1}\frac{1}{2}\left(B^2(t_{i+1}^n)-B^2(t_i)\right)-\frac{1}{2}\sum_{i=0}^{n-1}\left(B(t_{i+1}^n)-B(t_i^n)\right)^2$$

$$=\frac{1}{2}B^2(T)-\frac{1}{2}B^2(0)-\frac{1}{2}\sum_{i=0}^{n-1}\left(B(t_{i+1}^n)-B(t_i^n)\right)^2.$$

and then

$$\begin{aligned}
&\int_0^T X^n(t)dB(t) \\
=&\sum_{i=0}^{n-1} \frac{1}{2}\big(B^2(t_{i+1}^n) - B^2(t_i)\big) - \frac{1}{2}\sum_{i=0}^{n-1}\big(B(t_{i+1}^n) - B(t_i^n)\big)^2 \\
=&\frac{1}{2}B^2(T) - \frac{1}{2}B^2(0) - \frac{1}{2}\sum_{i=0}^{n-1}\big(B(t_{i+1}^n) - B(t_i^n)\big)^2.
\end{aligned}$$

The last summation converges to the quadratic variation T of Brownian motion on $[0,T]$ in probability. So, $\int_0^T X^n(t)dB(t)$ converges in probability, and the limit is

$$\int_0^T B(t)dB(t) = \lim \int_0^T X^n(t)dB(t) = \frac{1}{2}B^2(T) - \frac{1}{2}T.$$

Itô's Integral for square-integrable adapted processes

Theorem

Let $X(t)$ be an adapted process such that

$$\int_0^T \mathsf{E}[X^2(t)]dt < \infty. \tag{2.2.3}$$

Then Itô integral $\int_0^T X(t)dB(t)$ is defined and satisfied Properties (i)-(vii).

Proof. For an adapted process $X(t)$ there exists a sequence of simple process $X^n(t)$ such that

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |X(t) - X^n(t)|^2 dt = 0.$$

Proof. For an adapted process $X(t)$ there exists a sequence of simple process $X^n(t)$ such that

$$\lim_{n\to\infty} \mathsf{E}\int_0^T |X(t)-X^n(t)|^2 dt = 0.$$

Indeed, for a continuous process X stratifying (2.2.3), such $X^n(t)$ can be taken as

$$X(0)+\sum_{k=0}^{2^n-1} X(\frac{k}{2^n}T) I_{(kT/2^n,(k+1)T/2^n]}(t).$$

If X is not continuous, the construction of approximating processes is more involved (See the Lemma below).

The Itô integrals for simple processes $X^n(t)$ is defined by (2.2.2). By the Isometry property

$$\mathsf{E}\left[\int_0^T X^n(t)dB(t) - \int_0^T X^m(t)dB(t)\right]^2$$
$$= \int_0^T \mathsf{E}[X^n(t) - X^m(t)]^2 dt \to 0 \quad n, m \to \infty.$$

The Itô integrals for simple processes $X^n(t)$ is defined by (2.2.2). By the Isometry property

$$\mathsf{E}\left[\int_0^T X^n(t)dB(t) - \int_0^T X^m(t)dB(t)\right]^2$$
$$= \int_0^T \mathsf{E}[X^n(t) - X^m(t)]^2 dt \to 0 \quad n, m \to \infty.$$

That is, the Itô integrals for simple processes $X^n(t)$ form a Cauchy sequence in L_2. By the completeness of L_2, $\int_0^T X^n(t)dB(t)$ converges to a limit in L_2. We denote the limit of $\int_0^T X^n(t)dB(t)$ by Z.

If $Y^n(t)$ is also a simple processes satisfying

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |(X(t) - Y^n(t)|^2 dt = 0.$$

Then

$$\lim_{n\to\infty} \mathsf{E} \int_0^T |(X^n(t) - Y^n(t)|^2 dt = 0.$$

Using the Isometry property,

$$\mathsf{E}\left[\int_0^T Y^n(t)dB(t) - \int_0^T X^n(t)dB(t)\right]^2$$
$$= \int_0^T \mathsf{E}[Y^n(t) - X^n(t)]^2 dt \to 0 \quad n, m \to \infty.$$

So the limit $\int_0^T X^n(t)dB(t)$ does not depend on the choice of the approximating sequence.

The limit of $\int_0^T X^n(t)dB(t)$ is called $\int_0^T X(t)dB(t)$.

The limit of $\int_0^T X^n(t)dB(t)$ is called $\int_0^T X(t)dB(t)$.

Properties (i)-(v) and (vii) are easy to be verified by taking limits since the Itô integral of the simple processes satisfies Property (i)-(vii).

For Property (vi), we want to show that

$$[I,I](T)=\int_0^T X^2(u)du.$$

Let $\Pi=\{s_0,s_1,\ldots,s_m\}$ be a partition of $[0,T]$, and $\|\Pi\|$ be the largest length of the subinterval. Denote $Q(\Pi)$ and $Q_n(\Pi)$ be the sample quadratic variation of $I(t)=:\int_0^t X(u)dB(u)$ and $I_n(t)=:\int_0^t X^n(u)dB(u)$ over $[0,T]$ respectively.

For Property (vi), we want to show that

$$[I, I](T) = \int_0^T X^2(u)du.$$

Let $\Pi = \{s_0, s_1, \ldots, s_m\}$ be a partition of $[0, T]$, and $\|\Pi\|$ be the largest length of the subinterval. Denote $Q(\Pi)$ and $Q_n(\Pi)$ be the sample quadratic variation of $I(t) =: \int_0^t X(u)dB(u)$ and $I_n(t) =: \int_0^t X^n(u)dB(u)$ over $[0, T]$ respectively.

Notice $[I_n, I_n](T) = \int_0^T (X^n(u))^2 du \to \int_0^T (X(u))^2 du$ in probability. It suffices to show that

$$\sup_{\Pi} \mathsf{E}|Q_n(\Pi) - Q(\Pi)| \to 0 \quad n \to \infty.$$

Using the inequality

$$b^2 - a^2 = (b-a)^2 + 2a(b-a) \le (b-a)^2 + \epsilon^{-2}(b-a)^2 + \epsilon^2 a^2,$$

we obtain

$$\mathsf{E}\left|\left(\int_{s_i}^{s_{i+1}} X^n(u)dB(u)\right)^2 - \left(\int_{s_i}^{s_{i+1}} X(u)dB(u)\right)^2\right|$$

$$\le (1+\epsilon^{-2})\mathsf{E}\left(\int_{s_i}^{s_{i+1}} (X^n(u) - X(u))dB(u)\right)^2$$

$$+ \epsilon^2 \mathsf{E}\left(\int_{s_i}^{s_{i+1}} X(u)dB(u)\right)^2$$

$$\le (1+\epsilon^{-2}) \int_{s_i}^{s_{i+1}} \mathsf{E}\left[(X^n(u) - X(u))^2\right] du + \epsilon^2 \int_{s_i}^{s_{i+1}} \mathsf{E}\left[X^2(u)\right] du.$$

Adding up all these pieces, we obtain

$$\mathsf{E}|Q_n(\Pi) - Q(\Pi)| \leq (1 + \epsilon^{-2}) \int_0^T \mathsf{E}\left[(X^n(u) - X(u))^2\right] du + \epsilon^2 \int_0^T \mathsf{E}\left[X^2(u)\right] du.$$

The proof is now completed.

Lemma

Let $X(t)$ be an adapted process such that

$$\int_0^T \mathsf{E}[X^2(t)]dt < \infty. \tag{2.2.4}$$

Then there exists a sequence of simple process $\{X_n(t)\}$ such that

$$\int_0^T \mathsf{E}\left[\left(X_n(t) - X(t)\right)^2\right] dt \to 0. \tag{2.2.5}$$

Let $\mathscr{L}_2$ be a class of of adapted processes with (2.2.4). We complete the proof via three steps.

Step 1. Let $X \in \mathscr{L}_2$ be bounded and $X(\cdot, \omega)$ continuous for each ω. Then there exists a sequence of simple process $\{X_n(t)\}$ such that

$$\int_0^T \mathsf{E}\left[\left(X_n(t) - X(t)\right)^2\right] dt \to 0.$$

Let $\mathscr{L}_2$ be a class of of adapted processes with (2.2.4). We complete the proof via three steps.

Step 1. Let $X \in \mathscr{L}_2$ be bounded and $X(\cdot, \omega)$ continuous for each ω. Then there exists a sequence of simple process $\{X_n(t)\}$ such that

$$\int_0^T \mathsf{E}\left[\left(X_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof Take

$$X_n(t) = X(0) + \sum_{k=0}^{2^n-1} X(\frac{k}{2^n}T) I_{(kT/2^n,(k+1)T/2^n]}(t).$$

Step 2. Let $X \in \mathscr{L}_2$ be bounded. Then there exist bounded functions $Y_n \in \mathscr{L}_2$ such that $Y_n(\cdot, \omega)$ is continuous for all ω and n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Step 2. Let $X \in \mathscr{L}_2$ be bounded. Then there exist bounded functions $Y_n \in \mathscr{L}_2$ such that $Y_n(\cdot, \omega)$ is continuous for all ω and n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof. Suppose $|X(t, \omega)| \le M$. For each n, let ψ_n be a non-negative continuous real function such that

1. $\psi_n(x) = 0$ for $x \le -\frac{1}{n}$ and $x \ge 0$,
2. $\int_{-\infty}^{\infty} \psi_n(x) dx = 1$.

Define

$$Y_n(t,\omega) = \int_0^t \psi_n(s-t)X(s,\omega)ds.$$

Then $Y_n(\cdot,\omega)$ is continuous for all ω and $|Y_n(t,\omega)| \le M$.

Since $X \in \mathscr{L}_2$ we see that $Y_n(t,\omega)$ is $\mathscr{F}_t$-measurable for all t.

Moreover,

$$\int_0^T \big(Y_n(t,\omega) - X(t,\omega)\big)^2 dt \to 0, \quad \text{for each } \omega,$$

since $\{\psi_n\}$ constitutes an approximate identity.

Define

$$Y_n(t,\omega)=\int_0^t \psi_n(s-t)X(s,\omega)ds.$$

Then $Y_n(\cdot,\omega)$ is continuous for all ω and $|Y_n(t,\omega)|\le M$.

Since $X\in\mathscr{L}_2$ we see that $Y_n(t,\omega)$ is $\mathscr{F}_t$-measurable for all t.

Moreover,

$$\int_0^T \big(Y_n(t,\omega)-X(t,\omega)\big)^2 dt \to 0, \quad \text{for each } \omega,$$

since $\{\psi_n\}$ constitutes an approximate identity. So

$$\int_0^T \mathsf{E}\left[\big(Y_n(t)-X(t)\big)^2\right] dt \to 0$$

by bounded convergence.

Step 3. Let $X \in \mathscr{L}_2$. Then there exists a sequence $\{Y_n\} \subset \mathscr{L}_2$ such that Y_n is bounded for each n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Step 3. Let $X \in \mathscr{L}_2$. Then there exists a sequence $\{Y_n\} \subset \mathscr{L}_2$ such that Y_n is bounded for each n, and

$$\int_0^T \mathsf{E}\left[\left(Y_n(t) - X(t)\right)^2\right] dt \to 0.$$

Proof. Put

$$Y_n(t,\omega) = \begin{cases} -n & \text{if } X(t,\omega) < -n \\ X(t,\omega) & \text{if } -n \le X(t,\omega) \le n \\ n & \text{if } X(t,\omega) > n. \end{cases}$$

The the conclusion follows by dominated convergence.

Further properties of Itô's integral

Let $X(t), Y(t) \in \mathscr{L}_2$, and σ, τ be stopping times such that $\tau \le \sigma$. Then for any $0 \le t \le T$,

(viii)

$$\mathsf{E}\left[\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\Big|\mathscr{F}_{t\wedge\tau}\right] = 0,$$

$$\mathsf{E}\left[\left(\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\right)^2\Big|\mathscr{F}_{t\wedge\tau}\right] = \int_{t\wedge\tau}^{t\wedge\sigma} \mathsf{E}\left[X^2(u)\big|\mathscr{F}_{t\wedge\tau}\right] du.$$

Proof. Write $I_X(t) = \int_0^t X(t)dB(t)$, $M_X(t) = I_X^2(t) - [I_X, I_X](t)$. (viii) is equivalent to

$$\mathsf{E}\left[I_X(t \wedge \sigma)\middle|\mathscr{F}_{t\wedge\tau}\right] = I_X(t \wedge \tau),$$

$$\mathsf{E}\left[M_X(t \wedge \sigma)\middle|\mathscr{F}_{t\wedge\tau}\right] = M_X(t \wedge \tau).$$

Note $I_X(t)$ and $M_X(t)$ are both martingale. The results follows from the optional stopping time theorem (see the lemma below).

Lemma

(Optional stopping time theorem) If $M(t)$ is a right-continuous martingale with filtration $\{\mathscr{F}_t\}$ and τ, σ are bounded stopping times such that $P(\tau \le \sigma) = 1$. Then

$$\mathsf{E}[M(\sigma)|\mathscr{F}_\tau] = M(\tau).$$

(iv) For all $s \le t$,

$$\mathsf{E}\left[\left(\int_s^t X(u)dB(u)\right)\left(\int_s^t Y(u)dB(u)\right)\Big|\mathscr{F}_s\right]$$
$$= \int_s^t \mathsf{E}\left[X(u)Y(u))\big|\mathscr{F}_s\right] du,$$

$$\mathsf{E}\left[\left(\int_{t\wedge\tau}^{t\wedge\sigma} X(u)dB(u)\right)\left(\int_{t\wedge\tau}^{t\wedge\sigma} Y(u)dB(u)\right)\Big|\mathscr{F}_{t\wedge\tau}\right]$$
$$= \int_{t\wedge\tau}^{t\wedge\sigma} \mathsf{E}\left[X(u)Y(u)\big|\mathscr{F}_{t\wedge\tau}\right] du.$$

Proof. Write $M_{X,Y}(t) = I_X(t)I_Y(t) - [I_X, I_Y](t)$. Note $\mathsf{E}[I_X(t) - I_X(s)|\mathscr{F}_s] = 0$, $\mathsf{E}[I_Y(t) - I_Y(s)|\mathscr{F}_s] = 0$. The first property in (iv) is equivalent to

$$\mathsf{E}\left[M_{X,Y}(t)\middle|\mathscr{F}_s\right] = M_{X,Y}(s),$$

which is due to the fact that

Proof. Write $M_{X,Y}(t) = I_X(t)I_Y(t) - [I_X, I_Y](t)$. Note $\mathsf{E}[I_X(t) - I_X(s)|\mathscr{F}_s] = 0$, $\mathsf{E}[I_Y(t) - I_Y(s)|\mathscr{F}_s] = 0$. The first property in (iv) is equivalent to

$$\mathsf{E}\left[M_{X,Y}(t)\middle|\mathscr{F}_s\right] = M_{X,Y}(s),$$

which is due to the fact that

$$\begin{aligned} 2M_{X,Y}(t) =& I_{X+Y}^2(t) - [I_{X+Y}, I_{X+Y}](t) \\ & - \left(I_X^2(t) - [I_X, I_X](t)\right) - \left(I_Y^2(t) - [I_Y, I_Y](t)\right) \end{aligned}$$

is a martingale.

By the optional stopping time theorem, we have

$$\mathsf{E}\left[M_{X,Y}(t\wedge\sigma)\Big|\mathscr{F}_{s\wedge\tau}\right]=M_{X,Y}(s\wedge\tau),$$

which is equivalent to the second property in (iv).

(iiv) If σ is a stopping time, then

$$\int_0^{t\wedge\sigma} X(u)dB(u) = \int_0^t X'(u)dB(u), \quad \forall \ t \geq 0,$$

where $X'(t,\omega) = X(t,\omega)I_{\{\sigma(\omega)\geq t\}}$.

Proof. If Property (iiv) is checked for simple processes and then the general case can be proved by approximating X with simple processes. The following proof can be refered to N. Ikeda and Watanabe (1889), *Stochastic Differential Equations and Diffusion Process*, page 50, North-Holland Publishing Company.

Suppose X is a simple process,

$$X(t) = \xi_0 I_0(t) + \sum_i \xi_i I_{(t_i, t_{i+1}]}(t).$$

Let $\{s_j^{(n)}\}$ be a refinement of subdivisions $\{t_i\}$. Suppose X has the expression

$$X(t,\omega) = \xi_0 I_0(t) + \sum_j \xi_j^{(n)} I_{(s_j^{(n)}, s_{j+1}^{(n)}]}(t).$$

Define

$$\sigma^n(\omega) = s_{j+1}^{(n)} \quad \text{if} \quad \sigma(\omega) \in (s_j^{(n)}, s_{j+1}^{(n)}].$$

It is easy to see that σ^n is an $\mathscr{F}_t$-stopping time for each $n = 1, 2, \ldots$ and $\sigma^n \downarrow \sigma$ as $n \to \infty$. If $s \in (s_j^{(n)}, s_{j+1}^{(n)}]$, then $I_{\{\sigma^n \geq s\}} = I_{\{\sigma > s_j^{(n)}\}}$ and therefore, if we set $X_n'(s,\omega) = X(s,\omega) I_{\{\sigma^n(\omega) \geq s\}}$, then

$$\begin{aligned} X_n'(s,\omega) =& \xi_0 I_0(s) + \sum_j \xi_j^{(n)} I_{\{\sigma^n \geq s\}} I_{(s_j^{(n)}, s_{j+1}^{(n)}]}(s) \\ =& \xi_0 I_0(s) + \sum_j \xi_j^{(n)} I_{\{\sigma > s_j^{(n)}\}} I_{(s_j^{(n)}, s_{j+1}^{(n)}]}(s) \end{aligned}$$

is also a simple process.

Clearly, for every $t > 0$,

$$\mathsf{E}\left[I(X_n')(t) - I(X')(t)\right]^2 = \mathsf{E}\left[\int_0^t \left(X_n'(u) - X'(u)\right)^2 du\right]$$

$$= \mathsf{E}\left[\int_0^t X^2(u) I_{\{\sigma^n \geq u > \sigma\}} du\right] \to 0,$$

as $n \to \infty$. Hence $I(X_n')(t) \to I(X')(t)$ in probability.

Also

$$I(X_n')(t) = \sum_j \xi_j^{(n)} I_{\{\sigma > s_j^{(n)}\}} [B(t \wedge s_{j+1}^{(n)}) - B(t \wedge s_j^{(n)})]$$

$$= \sum_j \xi_j^{(n)} I_{\{\sigma > s_j^{(n)}\}} [B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})]$$

(since $\sigma^n \geq s_{j+1}^{(n)}$ if $\sigma > s_j^{(n)}$)

$$= \sum_j \xi_j^{(n)} [B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})]$$

(since $\sigma^n \leq s_j^{(n)}$ if $\sigma \leq s_j^{(n)}$

and so $[B(t \wedge \sigma^n \wedge s_{j+1}^{(n)}) - B(t \wedge \sigma^n \wedge s_j^{(n)})] = 0$)

$$= \int_0^{t \wedge \sigma^n} X(u) dB(u) = I(X)(t \wedge \sigma_n).$$

Consequently,

$$I(X')(t) = \lim I(X'_n)(t) = \lim I(X)(t \wedge \sigma^n) = I(X)(t \wedge \sigma),$$

by the continuity of $I(X)(t)$.

Itô's Integral for predictable processes

The approach of definiting of the Itô integral by approximation can be carried out for the class of predictable process $X(t)$, $0 \leq t \leq T$, satisfying the condition

$$\int_0^T X^2(t)dt < \infty \quad a.s. \tag{2.2.6}$$

Recall that a process X is adapted if for any t, the value $X(t)$ can depend on the past values of $B(s)$ for $s \leq t$, but not on future values of $B(s)$ for $s > t$.

Intuitively, an adapted process is predictable if for any t, the value $X(t)$ is determined by the values of values of $B(s)$ for $s < t$. X is predictable if it is

1. a left-continuous adaptive process, for example, a simple function, that is an adapted left-continues step function,
2. a limit (almost sure, in probability) of left-continuous adapted processes,
3. a Borel-measurable function of a predictable process.

For a predictable process X satisfying (2.2.6), one can choose simple processes X^n convergent to X in sense that

$$\int_0^T (X^n(t) - X(t))^2 dt \to 0 \text{ in probability.}$$

The sequence of Itô integrals $\int_0^T X^n(t)dB(t)$ is a Cauchy sequence in probability. It converges to a limit $\int_0^T X(t)dB(t)$. This limit does not depend on the choice of the sequence of simple processes. It is defined as the integral of X.

Theorem

The Itô integral of a predicable process X satisfying the condition (2.2.6) exists and it is a limit in probability of a sequence of corresponding Itô integrals of simple processes. Further, the integral $I(t) = \int_0^t X(u)dB(u)$ is a locally square integrable martingale, i.e., there exists a sequence of stopping times σ_n such that $\sigma_n < \infty$, $\sigma_n \uparrow \infty$ and $I_n(t) =: I(t \wedge \sigma_n)$ is a square integrable martingale for each n.

Proof. Let $\sigma_n(\omega) = \inf\{t : \int_0^t X^2(u,\omega)du \ge n\} \wedge n$. Then σ_n is a sequence of stopping times such that $\sigma_n \uparrow \infty$ a.s.. Set $X_n(t,\omega) = X(t,\omega) I_{\{\sigma_n(\omega)\ge t\}}$. Clearly

$$\int_0^\infty X_n^2(t,\omega)dt = \int_0^{\sigma_n} X^2(t,\omega)dt \le n,$$

and hence

$$\int_0^\infty \mathsf{E}[X_n^2(t)]dt \le n, \quad n = 1, 2, \ldots .$$

So, $I(X_n)(t) =: \int_0^t X_n(u)dB(u)$ is well defined and a square integrable martingale.

Notice $X_m(t,\omega) = X_n(t,\omega) I_{\{\sigma_m(\omega)\ge t\}}$ for $m < n$, by Property (iiv), we have for each $m < n$ that

$$I(X_n)(t \wedge \sigma_m) = I(X_m)(t).$$

Notice $X_m(t,\omega) = X_n(t,\omega) I_{\{\sigma_m(\omega)\ge t\}}$ for $m < n$, by Property (iiv), we have for each $m < n$ that

$$I(X_n)(t \wedge \sigma_m) = I(X_m)(t).$$

Consequently if we define $I(X)(t) = \int_0^t X(u)dB(u)$ by

$$I(X)(t) = I(X_n)(t) \ \text{ for } \ t \le \sigma_n,$$

then this definition is well-defined and determines a continuous process $I(X)$ such that

$$I(X)(t \wedge \sigma_n) = I(X_n)(t), \quad n = 1, 2, \ldots.$$

Therefore $I(X)$ is a local square integrable martingale.

Further, for each n, $I(X_n)$ is a limit in probability of integrals of a sequence of simple processes, and so is $I(X)$.

Further, for each n, $I(X_n)$ is a limit in probability of integrals of a sequence of simple processes, and so is $I(X)$.

Finally, it can be shown that if

$$\int_0^T (Y_n(t) - X(t))^2 dt \to 0 \text{ in probability,}$$

and Y_n is a simple process, then

$$I(Y_n)(t) \to I(X)(t) \text{ in probability.}$$

The proof is similar to the following theorem and omitted here.

Theorem

Suppose predicable processes X_n and X satisfy the condition (2.2.6) and

$$\int_0^T (X_n(t) - X(t))^2 dt \to 0 \quad \text{in probability.}$$

Then

$$\sup_{0\le t\le T} \left| \int_0^t X_n(u)dB(u) - \int_0^t X(u)dB(u) \right| \to 0 \quad \text{in probability.}$$

Proof. The result follows from the following inequality immediately.

$$\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X(t)dB(t)\right|\ge \epsilon\right)\le \mathsf{P}\left(\int_0^T X^2(t)dt\ge\delta\right)+\frac{\delta}{\epsilon^2}. \tag{2.2.7}$$

Proof. The result follows from the following inequality immediately.

$$P\left(\sup_{0\le t\le T}\left|\int_0^t X(t)dB(t)\right|\ge \epsilon\right)\le P\left(\int_0^T X^2(t)dt\ge\delta\right)+\frac{\delta}{\epsilon^2}. \tag{2.2.7}$$

For (2.2.7), we let

$$X_\delta(s)=X(s)I\{s:\int_0^s X^2(u)du<\delta\}\in\mathscr{L}_2.$$

Then $\int_0^t X_\delta(s)dB(s)$ is a square integrable martingale.

According to the Doob inequality,

$$\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X_\delta(t)dB(t)\right|\ge\epsilon\right)\le\frac{\mathsf{E}\left(\int_0^T X_\delta(t)dB(t)\right)^2}{\epsilon^2}$$

$$\le\frac{\mathsf{E}\left(\int_0^T X_\delta^2(t)dt\right)}{\epsilon^2}\le\frac{\delta}{\epsilon^2}.$$

It follows that

$$
\begin{aligned}
& \mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X(u)dB(u)\right|\ge\epsilon\right)\\
\le & \mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t \left(X(u)-X_\delta(u)\right)dB(u)\right|\ne 0\right)\\
& +\mathsf{P}\left(\sup_{0\le t\le T}\left|\int_0^t X_\delta(u)dB(u)\right|\ge\epsilon\right)\\
\le & \mathsf{P}\left(\int_0^T X^2(t)dt\ge\delta\right)+\frac{\delta}{\epsilon^2}.
\end{aligned}
$$

Remark The integral process $I(t) = \int_0^t X(u)dB(u)$ can be chosen to be a continuous process in sense that

$$\mathsf{P}\left(\{\omega : I(t,\omega) \text{ is continuous in } [0,T]\}\right) = 1.$$

Remark The integral process $I(t)=\int_0^t X(u)dB(u)$ can be chosen to be a continuous process in sense that

$$\mathsf{P}\left(\{\omega : I(t,\omega) \text{ is continuous in } [0,T]\}\right)=1.$$

In fact,

$$\sup_{0\le t\le T}\left|\int_0^t X_n(t)dB(t)-\int_0^t X_m(t)dB(t)\right|\to 0 \quad \text{in probability.}$$

Remark The integral process $I(t)=\int_0^t X(u)dB(u)$ can be chosen to be a continuous process in sense that

$$\mathsf{P}\left(\{\omega : I(t,\omega) \text{ is continuous in } [0,T]\}\right)=1.$$

In fact,

$$\sup_{0\le t\le T}\left|\int_0^t X_n(t)dB(t)-\int_0^t X_m(t)dB(t)\right|\to 0 \quad \text{in probability.}$$

So, there is a process $Z(t)$ such that

$$\sup_{0\le t\le T}\left|\int_0^t X_n(t)dB(t)-Z(t)\right|\to 0 \quad \text{in probability.}$$

So, there is a subsequence n_k such that

$$\sup_{0\le t\le T} \left| \int_0^t X_{n_k}(u)dB(u) - Z(t) \right| \to 0 \ \ a.s.$$

Note that $X_{n_k}(t)$ is a simple process. $\int_0^t X_{n_k}(u)dB(u)$ is a continuous process. So, $Z(t)$ is almost surely a continuous process.

Corollary

If X is a continuous adapted process then the Itô integral $\int_0^T X(t)dB(t)$ exists. Further, if $\{t_i^n\}$ is a partition of the interval $[0,T]$,

$$0 = t_0^n < t_1^n < \ldots < t_n^n = T,$$

with $\delta_n = \max_{1\le i\le n}(t_i^n - t_{i-1}^n) \to 0$, then

$$\int_0^T X(t)dB(t) = \lim_{n\to\infty}\sum_{i=0}^{n-1} X(t_i^n)\left[B(t_{i+1}^n)-B(t_i^n)\right] \quad \textit{in probability.}$$

Proof. A continuous adapted process X is a predictable process and $\int_0^T X^2(t)dt < \infty$ by the continuity. So, the Itô integral exists.

Proof. A continuous adapted process X is a predictable process and $\int_0^T X^2(t)dt < \infty$ by the continuity. So, the Itô integral exists. Let

$$X^n(t) = X(0)I_0(t) + \sum_{i=0}^{n-1} X(t_i^n) I_{(t_i^n, t_{i+1}^n]}(t)$$

for the partition $\{t_i^n\}$. Then X^n's are a simple processes and for any t, $X^n(t) \to X(t)$ as $n \to \infty$ by the continuity.

Proof. A continuous adapted process X is a predictable process and $\int_0^T X^2(t)dt < \infty$ by the continuity. So, the Itô integral exists. Let

$$X^n(t) = X(0)I_0(t) + \sum_{i=0}^{n-1} X(t_i^n)I_{(t_i^n, t_{i+1}^n]}(t)$$

for the partition $\{t_i^n\}$. Then X^n's are a simple processes and for any t, $X^n(t) \to X(t)$ as $n \to \infty$ by the continuity. For the simple process X^n, the Itô process is

$$\int_0^T X^n(t)dB(t) = \sum_{i=0}^{n-1} X(t_i^n)\big[B(t_{i+1}^n) - B(t_i^n)\big].$$

The result follows.

2.3 Itô's integral process and Stochastic differentials

Suppose X is a predictable process with $\int_0^T X^2(t)dt < \infty$ with probability one, then $\int_0^t X(u)dB(u)$ is well-defined for any $t \le T$. The process

$$Y(t) = \int_0^t X(u)dB(u)$$

is called the Itô integral process. It can be showed that $Y(t)$ has a version with continuous paths. In what follows it is always assumed $Y(t)$ is continuous. So Itô integral process $Y(t) = \int_0^t X(u)dB(u)$ is continuous and adapted.

The Itô integral can be written in the differential notation

$$dY(t) = X(t)dB(t).$$

The integral expression and the differential expression have the mean but with different notations.

It should be remembered that the differential expression does not mean that

$$\frac{Y(t+\Delta t)-Y(t)}{B(t+\Delta t)-B(t)} \to X(t) \quad \text{as } \Delta \to 0.$$

It is just a differential expression of the Itô integral, that is

$$Y(b)-Y(a)=\lim_n \sum_i X(t_i^n)\big[B(t_{i+1}^n)-B(t_i)\big],$$

when $X(t)$ is continuous .

If further $\int_0^T \mathsf{E}[X^2(t)]dt < \infty$, we have shown that $Y(t)$ having the following properties:

1. $Y(t)$ is a continuous square integrable martingale with $\mathsf{E}Y^2(t) = \int_0^t \mathsf{E}[X^2(u)]du$,
2. $[Y, Y](t) = \int_0^t X^2(u)du$,
3. $Y^2(t) - \int_0^t X^2(u)du$ is also a martingale.

When $\int_0^T \mathsf{E}[X^2(t)]dt = \infty$, Properties 1 and 3 may fail. Property 2 remains true.

Theorem

Quadratic variation of Itô integral process $Y(t) = \int_0^t X(u)dB(u)$ is given by

$$[Y, Y](t) = \int_0^t X^2(u)du.$$

Proof. Let $\sigma_m(\omega) = \inf\{t : \int_0^t X^2(u,\omega)du \ge m\} \wedge m$. Then σ_m is a sequence of stopping times such that $\sigma_m \uparrow \infty$ a.s.. Set $X_m(t,\omega) = X(t,\omega)I_{\{\sigma_m(\omega)\ge t\}}$. Clearly

$$\int_0^\infty X_m^2(t,\omega)dt = \int_0^{\sigma_m} X^2(t,\omega)dt \le m.$$

If define $Y_m(t) = \int_0^t X_m(u)dB(u)$, then

$$Y_m(t) = \int_0^{t\wedge\sigma_m} X(u)dB(u) = Y(t\wedge\sigma_m),$$

$$[Y_m, Y_m](t) = \int_0^t X_m^2(u)du = \int_0^{t\wedge\sigma_m} X^2(u)du.$$

So, on the event $\{\sigma_m > t\}$,

$$[Y_m, Y_m](t) = \int_0^t X^2(u)du.$$

On the other hand, on the event $\{\sigma_m > t\}$,

$$
\begin{aligned}
[Y_m, Y_m](t) &= \lim_n \sum_i \left| Y_m(t_{i+1}^n) - Y_m(t_i^n) \right|^2 \\
&= \lim_n \sum_i \left| Y(t_{i+1}^n \wedge \sigma_m) - Y(t_i^n \wedge \sigma_m) \right|^2 \\
&= \lim_n \sum_i \left| Y(t_{i+1}^n) - Y(t_i^n) \right|^2 = [Y, Y](t).
\end{aligned}
$$

The proof is completed.

With the same argument, one can show that If $\int_0^T X^2(t)dt < \infty$ a.s., then $Y(t) = \int_0^t X(u)dB(u)$ have the following properties:

1. $Y(t)$ is a continuous local square integrable martingale,
2. $Y^2(t) - \int_0^t X^2(u)du$ is also a local martingale.

Recall that the Quadratic variation of Y is defined by

$$[Y,Y](t) = \lim \sum_{i=0}^{n-1} \left| Y(t_{i+1}^n) - Y(t_i^n) \right|^2 \quad \text{in probability,}$$

when partition $\{t_i^n\}$ of $[0,t]$ become finer and finer. Loosely writing, this can be expressed as

$$\int_0^t \big(dY(u)\big)^2 = \int_0^t X^2(u)du$$

or in differential notation $dY(t)dY(t) = \big(dY(t)\big)^2 = X^2(t)dt$.

Especially,

$$dB(t)dB(t) = (dB(t))^2 = dt.$$

Hence

$$d[Y,Y](t) = dY(t)dY(t) = [X(t)dB(t)][X(t)dB(t)] = X^2(t)dt.$$

Quadratic covariation of Itô integrals

If $Y_1(t)$ and $Y_2(t)$ are Itô integrals of $X_1(t)$ and $X_2(t)$ with respect to the same Brownian motion $B(t)$, we define quadratic covariation of Y_1 and Y_2 on $[0,t]$ by

$$[Y_1, Y_2](t) = \frac{1}{2}\Big([Y_1+Y_2, Y_1+Y_2](t) - [Y_1, Y_1](t) - [Y_2, Y_2](t)\Big).$$

Then it follows that

$$[Y_1, Y_2](t) = \int_0^t X_1(s)X_2(s)ds. \tag{2.3.1}$$

It can be shown that quadratic covariation is given by the limit in probability of crossproducts of increments of processes

$$[Y_1, Y_2](t) = \lim \sum_{i=0}^{n-1} \left[Y_1(t_{i+1}^n) - Y_1(t_i^n)\right]\left[Y_2(t_{i+1}^n) - Y_2(t_i^n)\right], \tag{2.3.2}$$

when partition $\{t_i^n\}$ of $[0,t]$ become finer and finer.

(2.3.1) can be expressed as in differential notations

$$\begin{aligned} d[Y_1, Y_2](t) =& dY_1(t)dY_2(t) = X_1(t)dB(t)X_2(t)dB(t) \\ =& X_1(t)X_2(t)dt. \end{aligned}$$

In the limit in (2.3.2), if one of the function say Y_1 is of finite variation and the other is continuous, then it can be showed that the limit is zero.

Theorem

Suppose f is a real function of finite variation on interval $[a,b]$, and g is a continuous function on $[a,b]$. Then

$$[f,g]([a,b]) = \lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \left[f(t_{i+1}^n)-f(t_i^n)\right]\left[g(t_{i+1}^n)-g(t_i^n)\right] = 0,$$

where the limit is taken over all partitions $\Pi = \{t_i^n\}$ of $[a,b]$ with $\|\Pi\| = \max_i(t_{i+1}-t_i) \to 0$.

Proof. In fact, the summation is

$$\leq \max_i \left|g(t_{i+1}^n) - g(t_i^n)\right| \sum_i \left|f(t_{i+1}^n) - f(t_i^n)\right|$$

$$\leq \max_{|t-s|\leq\|\Pi\|} \left|g(t) - g(s)\right| V_f([a,b]) \to 0.$$

From this theorem, it follows that

$$\left[\int_0^s f(u)du, \int_0^s X(u)dB(u)\right](t) = 0.$$

This can be expressed as

$$\big(f(t)dt\big)\big(X(t)dB(t)\big) = 0.$$

Especially,

$$dtdB(t) = 0.$$

So, we arrive at

$$(dB(t))^2 = dt, \quad dtdB(t) = 0, \quad dtdt = 0. \tag{2.3.3}$$