

## 7. ITÔ INTEGRAL FROM ITÔ FORMULA

We will now continue the calculations from the previous section to ultimately “discover” the notion of the Itô integral.

## 7.1 Discovering the Itô integral.

Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function and let  $\{B_t: t \geq 0\}$  be standard Brownian motion. Suppose we wish to express the increment  $f(B_t) - f(B_0)$  using tools from differential and integral calculus similarly as one does in the theory of Riemann-Stieltjes integral. For this purpose we first note that, assuming  $f \in C^2(\mathbb{R})$ , Taylor’s Theorem along with a simple substitution give

$$\begin{aligned} f(y) &= f(x) + f'(x)(y-x) + \int_x^y f''(s)(y-s)dy \\ &= f(x) + f'(x)(y-x) + (y-x)^2 \int_0^1 f''((1-\theta)x + \theta y)(1-\theta)d\theta \end{aligned} \quad (7.1)$$

Given a partition  $\Pi = \{0 = t_0 < t_1 < \dots < t_n = t\}$  of  $[0, t]$ , we can thus apply this to each term of the telescopic representation  $f(B_t) - f(B_0)$  of to get

$$\begin{aligned} f(B_t) - f(B_0) &= \sum_{i=1}^n [f(B_{t_i}) - f(B_{t_{i-1}})] \\ &= \sum_{i=1}^n f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2} \sum_{i=1}^n f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2 \\ &\quad + \sum_{i=1}^n \left( \int_0^1 [f''((1-\theta)B_{t_{i-1}} + \theta B_{t_i}) - f''(B_{t_{i-1}})](1-\theta)d\theta \right) (B_{t_i} - B_{t_{i-1}})^2 \end{aligned} \quad (7.2)$$

We will now analyze the limit of the three sums on the right in the limit as the mesh of the partition tends to zero.

Leaving the first sum on the right alone until the very end, the second sum is treated by the trick we already employed in (6.22):

$$\begin{aligned} &\sum_{i=1}^n f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2 \\ &= \sum_{i=1}^n f''(B_{t_{i-1}})(t_i - t_{i-1}) + \sum_{i=1}^n f''(B_{t_{i-1}}) \left[ (B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right] \end{aligned} \quad (7.3)$$

Here we prove:

**Lemma 7.1** For all  $f \in C^2(\mathbb{R})$ , all  $t \geq 0$  and with  $\Pi$  denoting a generic partition of  $[0, t]$ ,

$$\sum_{i=1}^n f''(B_{t_{i-1}})^2 \left[ (B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right] \xrightarrow{\|\Pi\| \rightarrow 0} 0 \quad (7.4)$$

*Proof.* If  $f''$  is bounded, we can prove this by taking the second moment and noting that, due to square brackets being independent of their multiplies and each other, this reduces to the sum over  $\text{Var}((B_{t_i} - B_{t_{i-1}})^2)$ . To remove the assumption of boundedness of  $f''$ , we need to employ a truncation argument.

Fix  $\epsilon > 0$ . The continuity and local boundedness of Brownian paths implies that then there exists  $M > 0$  such that

$$P\left(\sup_{s \in [0, t]} |B_s| > M\right) < \epsilon. \quad (7.5)$$

Define  $C_i(\Pi) := \max\{f'(B_{t_{i-1}})^2, M\}$  and observe that on  $\{\max_{s \in [0, t]} |B_s| \leq M\}$  we have  $C_i(\Pi) = f'(B_{t_{i-1}})^2$ . Writing  $Q(\Pi)$  for the sum in the statement, we then get

$$P(|Q(\Pi)| > \epsilon) \leq \epsilon + P\left(\left|\sum_{i=1}^n C_i(\Pi) \left[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\right]\right| > \epsilon\right) \quad (7.6)$$

Since  $|C_i(\Pi)| \leq M$ , the second probability is bounded using Chebyshev's inequality by

$$\frac{1}{\epsilon^2} M^2 \sum_{i=1}^n \text{Var}((B_{t_i} - B_{t_{i-1}})^2)^2 \leq \frac{M^2}{\epsilon^2} 3t \|\Pi\|. \quad (7.7)$$

Taking  $\|\Pi\| \rightarrow 0$  followed by  $\epsilon \downarrow 0$  we get the claim.  $\square$

Since the first term on the right of (7.3) is a Riemann sum, we conclude:

**Corollary 7.2** For all  $f \in C^2(\mathbb{R})$ , all  $t \geq 0$  and with  $\Pi$  denoting a generic partition of  $[0, t]$ ,

$$\sum_{i=1}^n f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2 \xrightarrow{\|\Pi\| \rightarrow 0} \int_0^t f''(B_s) ds \quad (7.8)$$

For the error term in (7.2) we in turn get:

**Lemma 7.3** For all  $f \in C^2(\mathbb{R})$ , all  $t \geq 0$  and with  $\Pi$  denoting a generic partition of  $[0, t]$ ,

$$\sum_{i=1}^n \left( \int_0^1 \left[ f''((1-\theta)B_{t_{i-1}} + \theta B_{t_i}) - f''(B_{t_{i-1}}) \right] (1-\theta) d\theta \right) (B_{t_i} - B_{t_{i-1}})^2 \xrightarrow{\|\Pi\| \rightarrow 0} 0 \quad (7.9)$$

*Proof.* Denoting the oscillation of  $f$  on  $A$  over scales up to  $\delta$  by

$$\text{osc}_f(A, \delta) := \sup\{|f(x) - f(x')| : x, x' \in A \wedge |x - x'| < \delta\}, \quad (7.10)$$

the sum in the statement is bounded by

$$\frac{1}{2} \text{osc}_{f''} \left( B([0, t]), \text{osc}_B([0, t], \|\Pi\|) \right) V_t^{(2)}(B, \Pi) \quad (7.11)$$

By continuity of  $B$ , the inner oscillation tends to zero as  $\|\Pi\| \rightarrow 0$ . The outer oscillation then does as well since  $f''$  is continuous and  $B([0, t])$  is compact. Since  $V_t^{(2)}(B, \Pi)$  is bounded in probability, the whole quantity tends to zero in probability as  $\|\Pi\| \rightarrow 0$ .  $\square$

Returning to the first term, we can directly conclude:

**Corollary 7.4** For all  $h \in C^1(\mathbb{R})$  and all  $t \geq 0$ , there exists a random variable  $I_t(h)$  such that for  $\Pi$  denoting a generic partition of  $[0, t]$ ,

$$\sum_{i=1}^n h(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \xrightarrow[\|\Pi\| \rightarrow 0]{P} I_t(h) \quad (7.12)$$

Moreover, for  $f \in C^2(\mathbb{R})$  we then have

$$f(B_t) - f(B_0) = I_t(f') + \frac{1}{2} \int_0^t f''(B_s) ds \quad \text{a.s.} \quad (7.13)$$

*Proof.* The left-hand side of (7.2) is independent of  $\Pi$  and by the above reasoning, the second and the third sum converge to  $\frac{1}{2} \int_0^t f''(B_s)$  in probability. Hence also the first sum must converge to a random variable that satisfies (7.13). Replacing  $f$  by its antiderivative gives convergence in (7.12) for all  $h \in C^1(\mathbb{R})$ .  $\square$

## 7.2 Finessing the details.

The requirement that  $h$  be continuously differentiable for (7.12) to hold is actually redundant. We check this in:

**Lemma 7.5** The limit (7.12) exists for all  $h \in C(\mathbb{R})$ .

*Proof.* Let  $h \in C(\mathbb{R})$ . Given  $\epsilon > 0$ , let  $M > 0$  be such that (7.5) holds. The Weierstrass approximation theorem gives us  $h_\epsilon \in C^1(\mathbb{R})$  such that

$$\sup_{x \in [-M, M]} |h(x) - h_\epsilon(x)| < \epsilon \quad (7.14)$$

Denoting

$$I_t(h, \Pi) := \sum_{i=1}^n h(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \quad (7.15)$$

we then have

$$P\left(|I_t(h, \Pi) - I_t(h_\epsilon, \Pi)| > \delta\right) \leq \epsilon + P\left(\left|\sum_{i=1}^n \tilde{C}_i(\Pi)(B_{t_i} - B_{t_{i-1}})\right| > \delta\right) \quad (7.16)$$

where  $\tilde{C}_i(\Pi) := (h(B_{t_{i-1}}) \wedge \epsilon) \vee (-\epsilon)$ . Using  $|\tilde{C}_i(\Pi)| \leq \epsilon$ , the second probability is bounded using Chebyshev's inequality by  $\epsilon^2 t \delta^{-2}$  resulting in

$$P\left(|I_t(h, \Pi) - I_t(h_\epsilon, \Pi)| > \delta\right) \leq \epsilon(1 + t\epsilon\delta^{-2}) \quad (7.17)$$

Taking  $\|\Pi\| \rightarrow 0$  with the help of Fatou's lemma and the fact that  $I_t(h_\epsilon, \Pi) \rightarrow I_t(h)$  since  $h_\epsilon \in C^1$  translates this into

$$\lim_{r \downarrow 0} \sup_{\Pi: \|\Pi\| < r} P\left(|I_t(h, \Pi) - I_t(h_\epsilon)| > \delta\right) \leq \epsilon(1 + t\epsilon\delta^{-2}) \quad (7.18)$$

The triangle inequality then gives that, for all  $0 < \tilde{\epsilon} < \epsilon$ ,

$$P\left(|I_t(h_{\tilde{\epsilon}}) - I_t(h_{\epsilon})| > 2\delta\right) \leq 2\epsilon(1 + t\epsilon\delta^{-2}) \quad (7.19)$$

which by taking  $\epsilon \downarrow 0$  shows that  $I_t(h_{\epsilon})$  converges in probability to some  $I_t(h)$ . Using this fact in (7.18) then shows that  $I_t(h, \Pi) \rightarrow I_t(h)$  in probability as  $\|\Pi\| \rightarrow 0$ .  $\square$

Summarizing the above developments, we have thus proved two theorems:

**Theorem 7.6** (Itô integral) *Let  $h \in C(\mathbb{R})$ . For each  $t \geq 0$  there exists a random variable  $I_t(h)$  denoted also as  $\int_0^t h(B_s)dB_s$  such that, with  $\Pi$  denoting a generic partition of  $[0, t]$ ,*

$$\sum_{i=1}^n h(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \xrightarrow{\|\Pi\| \rightarrow 0} \int_0^t h(B_s)dB_s \quad (7.20)$$

**Theorem 7.7** (Itô formula) *For all  $f \in C^2$  and all  $t \geq 0$ ,*

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)dB_s + \frac{1}{2} \int_0^t f''(B_s)ds \quad \text{a.s.} \quad (7.21)$$

where the second term is the Itô integral from Theorem 7.6 and the second term is the ordinary Riemann integral.

These fact were first discovered by K. Itô who work on these in the dark days of wartime Japan. As was found out much later from some correspondence that sat in French Academy archives for decades, an independent discovery of these was made by W. Doebelin who perished on the French-German front in World War II. The integral  $\int_0^t h(B_s)dB_s$  is called the *Itô integral* or sometimes also the *stochastic integral*. This is justified by it being a limit of Riemann-Stieltjes sums.

The upshot of Itô's findings is twofold: First, while we write integral in (7.20) using the notation of the Riemann-Stieltjes integral, it actually does not fall under the umbrella of Riemann-Stieltjes integration. This is seen already in the simplest nontrivial case of function  $f(x) := x$  where the difference between the right-endpoint and left-endpoint Riemann-Stieltjes sums results in

$$\sum_{i=1}^n B_{t_i}(B_{t_i} - B_{t_{i-1}}) - \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) = \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2 \xrightarrow{\|\Pi\| \rightarrow 0} t \quad (7.22)$$

Another (closely related) fact is that a new set of calculus rules is needed for work with this integral as exemplified by the differential form

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt \quad (7.23)$$

The second integral on the right of (7.21) (as well as the second term on the right of (7.23)) is referred to as the *Itô correction*.

Since the stochastic integral is not a proper Riemann-Stieltjes integral, a natural question arises what happens when other rules are used inside the associated Riemann-Stieltjes sum. The following lemma will be relegated to a homework exercise:

**Lemma 7.8** Fix  $\theta \in [0, 1]$  and let  $\{B_t: t \in [0, \infty)\}$  be a standard Brownian motion. For any  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $t \geq 0$  and a partition  $\Pi = \{0 = t_0 < t_1 < \dots < t_n = t\}$  of  $[0, t]$  set

$$I_t^{(\theta)}(f, \Pi) := \sum_{i=1}^n f(B_{(1-\theta)t_{i-1} + \theta t_i})(B_{t_i} - B_{t_{i-1}}) \quad (7.24)$$

Assuming  $f \in C^1(\mathbb{R})$  with  $f, f'$  bounded, we then have

$$I_t^{(\theta)}(f, \Pi) \xrightarrow[\|\Pi\| \rightarrow 0]{P} \int_0^t f(B_s) dB_s + \theta \int_0^t f'(B_s) ds \quad (7.25)$$

Here the first integral on the right is as in (7.20).

For the “mid point rule”,  $\theta := 1/2$ , the limit object on the right is called the *Stratonovich integral*. This integral can be defined directly from the Itô integral as

$$\int_0^t f(B_s) \circ dB_s := \int_0^t f(B_s) dB_s + \frac{1}{2} \int_0^t f'(B_s) ds \quad (7.26)$$

The choice  $\theta = 1/2$  ensures that the Fundamental Theorem of Calculus is restored,

$$f(B_t) - f(B_0) = \int_0^t f(B_s) \circ dB_s, \quad (7.27)$$

thus making the Stratonovich integral attractive in certain situations (such as analysis on manifolds) where the Itô term wrecks havoc in calculations. The Stratonovich integral also arises naturally in solutions of stochastic differential equations obtained by approximating standard Brownian motion by smoothed-out versions thereof. (Indeed, the FTC holds in these and thus survives the limit as all approximations are taken away.) Still, the Stratonovich integral is not good for theory development as it requires that  $f$  be differentiable for it to be integrable. For this reason we will work mostly with the Itô integral in the sequel.

Further reading: Chapter 3 of Øksendal