8. Stochastic integral via Itô isometry

We will now move to define the stochastic integral in full scope of its validity. The main novelty is the extension from integrals of the form $\int_0^t f(B_s)\mathrm{d}B_s$ to integrals of the form $\int_0^t Y_s \mathrm{d}B_s$, where Y_s may depend on $\{B_u : u \leqslant s\}$ not just B_s alone. From the construction of the Paley-Zygmund integral we in turn draw the important idea to base the construction of the integral on L^2-isometry rather than pointwise convergence that underlies the Riemann or Lebesgue integration theory.

We start by a precise formulation of the phrase "Y_s may depend on $\{B_u : u \leqslant s\}$." This needs the following concept:

Definition 8.1 *Let $\{B_t : t \in [0,\infty)\}$ be a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$. A collection of σ-algebras $\{\mathcal{F}_t\}_{t\geqslant 0}$ is a* Brownian filtration *if*

(1) *$\{\mathcal{F}_t\}_{t\geqslant 0}$ is a filtration on $(\Omega, \mathcal{F})$, meaning that*

$$\forall 0 \leqslant s \leqslant t\colon \quad \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F} \tag{8.1}$$

(2) *B is adapted to $\{\mathcal{F}_t\}_{t\geqslant 0}$, meaning that*

$$\forall t \geqslant 0\colon \quad B_t \text{ is } \mathcal{F}_t\text{-measurable} \tag{8.2}$$

(3) *B is Markov with respect to $\{\mathcal{F}_t\}_{t\geqslant 0}$, meaning that*

$$\forall t \geqslant 0\colon \quad \sigma\bigl(B_{t+s} - B_t : s \geqslant 0\bigr) \perp\!\!\!\perp \mathcal{F}_t \tag{8.3}$$

Here we recall that σ-algebras $\mathcal{F}$ and $\mathcal{G}$ are said to be *independent*, with notation $\mathcal{F} \perp\!\!\!\perp \mathcal{G}$, if $\forall A \in \mathcal{F}\, \forall B \in \mathcal{G}\colon P(A \cap B) = P(A)P(B)$. We call (3) a Markov property because it ensures that a standard Brownian motion, if conditioned on $\mathcal{F}_t$ and reduced by the value of B_t, is a again a standard Brownian motion.

Our next task is the precise definition of "simple processes" which are those where we will easily agree on what the integral should evaluate to.

Definition 8.2 *Given a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ on a probability space $(\Omega, \mathcal{F}, P)$, a stochastic process $\{Y_t : t \in [0,\infty)\}$ on this space is said to be* simple *if there exist $n \in \mathbb{N}$, reals $0 =: t_0 < t_1 < \dots < t_n$ and random variables $Z_0, \dots, Z_n$ satisfying*

$$\forall i = 0, \dots, n\colon \quad Z_i \in L^\infty(\Omega, \mathcal{F}, P) \ \wedge \ Z_i \text{ is } \mathcal{F}_{t_{i-1}}\text{-measurable} \tag{8.4}$$

with the proviso $t_{-1} := 0$ such that

$$\forall t \geqslant 0\colon \quad Y_t = Z_0 1_{\{0\}}(t) + \sum_{i=1}^n Z_i 1_{(t_{i-1}, t_i]}(t) \tag{8.5}$$

We write $\mathcal{V}_0$ for the class of simple processes.

We remark that Z_0 will be irrelevant for the stochastic integral and so the restrictions imposed on it are done only for formal reasons (such has having Y adapted to the filtration). Calling the process "simple" is somewhat in conflict with the theory of the Lebesgue integral. Indeed, there the attribute "simple" is reserved to functions that are piecewise constant and measurable (and thus constant on measurable sets), while those constant on intervals are usually referred to as “step” functions. (Replacing simple functions by step functions is exactly what distinguishes Lebesgue integrability from Riemann integrability.)

The definition of a stochastic integral $\int_0^t Y_s \mathrm{d}B_s$ for Y simple is now fairly intuitive: multiply the value of the process on interval $(t_{i-1} - t_i]$ by the increment of the Brownian motion over this interval, and add these over all of the intervals contained in $[0,t]$ taking only the corresponding portion of the interval containing t. A problem is that the representation (8.5) of a simple process is not unique. We thus first note:

Lemma 8.3 *For $Y \in \mathcal{V}_0$ be given by* (8.5) *with a Brownian filtration associated with a Brownian motion $\{B_t \colon t \in [0,\infty)\}$. Then, for each $t \geqslant 0$, the expression on the right of*

$$\int_0^t Y_s \,\mathrm{d}B_s := \sum_{i=1}^n Z_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) \tag{8.6}$$

does not depend on the representation in (8.5). *Moreover, the integral is linear in the sense that, for all $t \geqslant 0$, all $Y, Y' \in \mathcal{V}_0$ and all $\alpha, \beta \in \mathbb{R}$,*

$$\int_0^t (\alpha Y_s + \beta Y'_s) \,\mathrm{d}B_s = \alpha \int_0^t Y_s \,\mathrm{d}B_s + \beta \int_0^t Y'_s \,\mathrm{d}B_s \tag{8.7}$$

Proof (idea). As for the Riemann integral, independence of representation is shown by finding a partition of $[0,t]$ that is common for any two ways to express Y as in (8.5) and then comparing the resulting expressions (8.6). The linearity (8.7) is checked similarly. We leave the details to the reader. □

Note that the definition (8.6) agrees with the quantity $I_t(f,\Pi)$ from (6.24) for Y_t given by (8.6) with $Z_i := f(B_{t_{i-1}})$ and Π determined by points $t_1, \dots, t_n = t$. The following lemma then extends Lemma 6.6 to our (more general) setting:

Lemma 8.4 (Itô isometry) *For all $t \geqslant 0$ and all $Y \in \mathcal{V}_0$,*

$$E\biggl[\Bigl(\int_0^t Y_s \,\mathrm{d}B_s\Bigr)^2\biggr] = E\biggl[\int_0^t Y_s^2 \mathrm{d}s\biggr], \tag{8.8}$$

where the integral on the right is in the Lebesgue (as well as Riemann) sense.

Proof. Writing the square of a sum as a double sum, the left-hand side of (8.8) equals

$$\sum_{i,j=1}^n E\Bigl(Z_i Z_j (B_{t_i \wedge t} - B_{t_{i-1} \wedge t})(B_{t_j \wedge t} - B_{t_{j-1} \wedge t})\Bigr) \tag{8.9}$$

Recall that Z_i is $\mathcal{F}_{t_{i-1}}$ measurable. This means that, for each $i < j$, the expectation above can be written as

$$E\Bigl(Z_i Z_j (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) E\bigl((B_{t_j \wedge t} - B_{t_{j-1} \wedge t}) \,\big|\, \mathcal{F}_{t_{j-1}}\bigr)\Bigr), \tag{8.10}$$

where we used that the terms Z_i, Z_j and $B_{t_i\wedge t}-B_{t_{i-1}\wedge t}$ are all $\mathcal{F}_{t_{j-1}}$-measurable. Since condition (3) in the definition of Brownian filtration implies

$$E\big((B_{t_j\wedge t}-B_{t_{j-1}\wedge t})\,\big|\,\mathcal{F}_{t_{j-1}}\big)\overset{\text{a.s.}}{=}E(B_{t_j\wedge t}-B_{t_{j-1}\wedge t})=0, \tag{8.11}$$

the expectation in (8.10) vanishes. A similar argument applies to the terms with $i>j$ which means that (8.9) equals

$$\sum_{i=1}^n E\Big(Z_i^2(B_{t_i\wedge t}-B_{t_{i-1}\wedge t})^2\Big) \tag{8.12}$$

Using again that Z_i is $\mathcal{F}_{t_{i-1}}$ measurable and that $B_{t_i\wedge t}-B_{t_{i-1}}$ is independent of $\mathcal{F}_{t_{i-1}}$, the expectation equals

$$E(Z_i^2)E\big((B_{t_i\wedge t}-B_{t_{i-1}\wedge t})^2\big)=E(Z_i^2)(t\wedge t_i-t\wedge t_{i-1}). \tag{8.13}$$

The left-hand side of (8.8) thus equals the expectation of

$$\sum_{i=1}^n Z_i^2(t_i\wedge t-t_{i-1}\wedge t). \tag{8.14}$$

This is now readily checked to equal $\int_0^t Y_s^2\mathrm{d}s$. □

In order to use the Itô isometry, we need to identify a suitable L^2-space of stochastic processes where the square-root of the right-hand side plays the role of L^2-norm. The minor problem is that the expression in (8.8) is for a fixed t, while we would like to work with processes on all of $[0,\infty)$. We thus put forward:

Definition 8.5 *Given a Brownian motion and a Brownian filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ on a probability space $(\Omega,\mathcal{F},P)$, we write $\mathcal{V}$ for the class of processes $\{Y_t\colon t\in[0,\infty)\}$ such that*

1. $(\omega,t)\mapsto Y_t(\omega)$ *is jointly measurable, meaning that, as a map $\Omega\times[0,\infty)\to\mathbb{R}$, it is measurable with respect to the product σ-algebra $\mathcal{F}\otimes\mathcal{B}([0,\infty))$,*
2. Y *is adapted, meaning $\forall t\geqslant 0\colon Y_t$ is $\mathcal{F}_t$-measurable,*
3. Y *is locally (jointly) square integrable, meaning*

$$\forall t>0\colon\quad \|Y\|_{L^2([0,t]\times\Omega)}:=\left(E\left[\int_0^t Y_s^2\mathrm{d}s\right]\right)^{1/2}<\infty. \tag{8.15}$$

Note that condition (1) implies (via Fubini-Tonelli) that $t\mapsto Y_t(\omega)$ is Borel function for each $\omega\in\Omega$. This will be quite useful in a number of arguments below. Condition (2) is in turn an abstract formulation of the fact that Y_t is independent of $B_{t+\cdot}-B_t$ which (as noted above) is key for the formulation of the Itô integral. We now observe:

Lemma 8.6 *$\mathcal{V}$ is a linear vector space with respect to pointwise addition and scalar multiplication. Moreoer, we have $\mathcal{V}_0\subseteq\mathcal{V}$; i.e., $\mathcal{V}$ contains all simple processes. In addition, defining*

$$\forall Y\in\mathcal{V}\colon\quad [\![Y]\!]:=\sum_{n\geqslant 1}2^{-n}\big(\|Y\|_{L^2([0,n]\times\Omega)}\wedge 1\big), \tag{8.16}$$

the map $Y, Y' \mapsto [\![Y - Y']\!]$ is a pseudometric on $\mathcal{V}$ such that for all $Y, Y' \in \mathcal{V}$,

$$[\![Y - Y']\!] = 0 \iff \lambda \otimes P\Big(\big\{(t,\omega) \in [0,\infty) \times \Omega\colon Y_t(\omega) = Y'_t(\omega)\big\}\Big) = 0 \tag{8.17}$$

where λ is the Lebesgue measure on $[0,\infty)$. The $\{\{Y' \in \mathcal{V}\colon [\![Y - Y']\!] = 0\}\colon Y \in \mathcal{V}\}$ of equivalence classes has the structure of a complete topological vector space.

We leave the (straightforward) proof of this lemma to homework exercise. The important consequence of Lemma 8.4 is then:

Corollary 8.7 *Denote by $\overline{\mathcal{V}_0}^{[\![\cdot]\!]}$ is the closure of $\mathcal{V}_0$ in $\mathcal{V}$ under the above topology. Then for each $t \geqslant 0$, the map $Y \mapsto \int_0^t Y_s \mathrm{d}B_s$ extends continuously to all $Y \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$. In particular, for all $t \geqslant 0$,* (8.8) *holds for all $Y \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$ and* (8.7) *holds a.s. for all $Y, Y' \in \mathcal{V}$ and all $\alpha, \beta \in \mathbb{R}$.*

Proof. Thanks to the underlying metric structure of $\mathcal{V}$ factored by the equivalence classes of processes for which the right-hand side (8.17) applies, it suffices to check that if $\{Y^{(n)}\}_{n\geqslant 1}$ is Cauchy in the sense that $[\![Y^{(n)} - Y^{(m)}]\!] \to 0$ as $n, m \to \infty$, then for each $t \geqslant 0$ the sequence of integrals $\{\int_0^t Y^{(n)} \mathrm{d}B_s\}_{n\geqslant 1}$ is Cauchy in L^2. This follows from (8.8) and (8.16). As the extension is based on L^2-limits, the isometry (8.8) remains in force. The additivity (8.7) is proved similarly. □

A natural question is now what processes otherthan the simple ones belong to the closure $\overline{\mathcal{V}_0}^{[\![\cdot]\!]}$ in $\mathcal{V}$. We will answer this in full completeness later, but for now we contend ourselves with:

Lemma 8.8 *Let $Y \in \mathcal{V}$ have left-continuous paths. Then $Y \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$.*

Proof. Suppose first that $Y \in \mathcal{V}$ is bounded, meaning $\exists c > 0\colon \sup_{t\geqslant 0} |Y_t| \leqslant c$ a.s., with left-continuous paths. Define, for each integer $n \geqslant 1$,

$$Y_t^{(n)} := Y_0 1_{\{0\}}(t) + \sum_{k=0}^{4^n} Y_{k2^{-n}} 1_{(k2^{-n},(k+1)2^{-n}]}(t) \tag{8.18}$$

Then $Y^{(n)} \in \mathcal{V}_0$ and, since $Y_t^{(n)} = Y_{2^n\lceil 2^n t\rceil - 2^{-n}}$, the left continuity shows

$$\forall t \geqslant 0\colon \quad Y_t^{(n)} \underset{n\to\infty}{\longrightarrow} Y_t \tag{8.19}$$

The Bounded Convergence Theorem now implies $\|Y^{(n)} - Y\|_{L^2([0,t]\times\Omega)} \to 0$ for each $t \geqslant 0$ and, consequently, $Y^{(n)} \to Y$ in $\mathcal{V}$. Hence $Y \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$.

Next assume that $Y \in \mathcal{V}$ is just left continuous and define, for each $M > 0$,

$$\widetilde{Y}_t^{(M)} := Y_T \wedge M \vee (-M). \tag{8.20}$$

Then $\widetilde{Y}^{(M)}$ is bounded and left-continuous and so $\widetilde{Y}^{(M)} \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$ by the previous argument. A calculation shows

$$\|\widetilde{Y}^{(M)} - \widetilde{Y}\|^2_{L^2([0,t]\times\Omega)} \leqslant E\int_0^t Y_s^2 1_{\{|Y_s|>m\}} \mathrm{d}s \tag{8.21}$$

which tends to zero as $M \to \infty$ by the Dominated Convergence Theorem. Hence we get $Y \in \overline{\mathcal{V}_0}^{[\![\cdot]\!]}$ as desired. $\square$

Lemma 8.8 implies that all continuous $Y \in \mathcal{V}$ are Itô integrable. In particular, this includes processes of the form $Y_s := f(B_s)$ where $f \in C(\mathbb{R})$ is — using Fubini-Tonelli and Definition 8.5(1) — such that $s \mapsto E[f(B_s)^2]$ is locally Lebesgue integrable. This alone is already an extension of Corollary 6.7 with more yet to come.

Since we already have a decent class of integrable processes, a question is: Can the integral be actually ever computed? As usual in other integration theories, this is not expected to apply to more than a handful of cases so we may as well start looking at particular examples. One strategy that works is to convert the stochastic integral to an ordinary Riemann or Stieltjes integral as in:

Lemma 8.9 *Let $f\colon [0,t] \to \mathbb{R}$ be of bounded variation (meaning $\sup_\Pi V_t^{(1)}(f,\Pi) < \infty$). Prove that*

$$\int_0^t f(s)\,\mathrm{d}B_s = f(t)B_t - f(0)B_0 - \int_0^t B_s\,\mathrm{d}f(s) \quad \text{a.s.}$$

where the latter is a Stieltjes integral. In fact, the same holds even if f is of bounded p-variation (meaning $\sup_\Pi V_t^{(p)}(f,\Pi) < \infty$) for some $p \in [1,2)$.

We leave the proof of this statement to homework. Another way to proceed is to invoke the Itô formula that we already proved in Lemma 6.5. Indeed, (6.26) implies that, for all $f \in C^1(\mathbb{R})$ we have

$$\int_0^t f(B_s)\mathrm{d}B_s = F(B_s) - F(B_0) - \frac{1}{2}\int_0^t f'(B_s)\mathrm{d}s \tag{8.22}$$

where $F\colon \mathbb{R} \to \mathbb{R}$ is any antiderivative of f, i.e., $F \in C^1(\mathbb{R})$ such that $F' = f$. For the special case $f(x) = x$, this yields

$$\int_0^t B_s\mathrm{d}B_s = \frac{1}{2}(B_t^2 - t) \tag{8.23}$$

Similarly we can compute stochastic integrals of $\int_0^t B_s^n\mathrm{d}B_s$, although the computation will get the harder the larger $n \geqslant 1$ gets. A unified approach to these integrals will be presented later.

Further reading: Section 3.1 in Øksendal