# MTH 9831. LECTURE 3

ELENA KOSYGINA

ABSTRACT. In this lecture we construct Itô integral.

1. Quadratic variation and cross-variation. The meaning of "multiplication rules" $dtdt = 0$, $dB(t)dB(t) = dt$, $dB(t)dt = 0$.
2. "Bare hands" approach to stochastic integral: the calculation of $\int_0^T B(t)\,dB(t)$ from the definition.
3. Itô integral for elementary integrands and its properties.
4. Extension of Itô integral to general integrands.

## 1. Cross-variation and quadratic variation.

Let $f, g : [0, T] \to \mathbb{R}$, $\Pi$ be a partition of $[0, T]$: $0 = t_0 < t_1 < \cdots < t_n = T$, and

$$\|\Pi\| := \max_{i \in \{1,2,\ldots,n\}} (t_i - t_{i-1})$$

be the length of the largest interval in the partition $\Pi$.

**Definition 1.1.** *The cross-variation of $f$ and $g$ on $[0, T]$ is defined by*

$$[f, g]_T := \lim_{\|\Pi\|\to 0} \sum_{i=1}^{n} (f(t_i) - f(t_{i-1}))(g(t_i) - g(t_{i-1}))$$

*if this limit exists. The quadratic variation $[f]_T$ of $f$ on $[0, T]$ is defined by $[f]_T := [f, f]_T$ when $[f, f]_T$ exists.*

Let us mention a useful polarization identity[1]:

$$2[f, g]_T = [f + g]_T - [f]_T - [g]_T. \tag{1.1}$$

The proof is obtained by using the identity $2xy = (x + y)^2 - x^2 - y^2$ in each term of the sum defining the left-hand side and taking limits. The details are left as an exercise.

**Example 1.2.** Let $f(t) = g(t) = t$. Then the cross-variation of $f$ and $g$ on $[0, T]$ is the same as the quadratic variation of either of them and is equal to 0. Indeed,

$$\sum_{i=1}^{n} (t_i - t_{i-1})^2 \le \|\Pi\| \sum_{i=1}^{n} (t_i - t_{i-1}) = \|\Pi\| T \to 0 \quad \text{as } \|\Pi\| \to 0.$$

This fact gives a precise meaning to the informal expression $dtdt = 0$, i.e. the last simply means that for $f(t) = t$ the quadratic variation of $f$ on $[0, T]$ is equal to 0 for all $T > 0$.

**Example 1.3.** Let $f$ and $g$ be differentiable functions on $[0, T]$ and such that their derivatives are square integrable (Riemann integral), i.e.

$$\int_0^T (f'(t))^2\,dt < \infty \quad \text{and} \quad \int_0^T (g'(t))^2\,dt < \infty.$$

Then $[f, g]_T = 0$.

[1] at least when any 3 of the 4 terms exist and finite.

By the mean value theorem we have

$$\sum_{i=1}^{n}(f(t_i)-f(t_{i-1}))(g(t_i)-g(t_{i-1}))=\sum_{i=1}^{n} f'(t_i^*)(t_i-t_{i-1})g'(\widetilde{t}_i)(t_i-t_{i-1}),$$

where $t_{i-1}\le t_i^*,\widetilde{t}_i\le t_i$, $i\in\{1,2,\dots,n\}$. By the Cauchy-Schwarz inequality[2], the absolute value of the last expression does not exceed

$$n\left(\frac{1}{n}\sum_{i=1}^{n}(f'(t_i^*))^2(t_i-t_{i-1})^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}(g'(\widetilde{t}_i))^2(t_i-t_{i-1})^2\right)^{1/2}$$
$$\le\|\Pi\|\left(\sum_{i=1}^{n}(f'(t_i^*))^2(t_i-t_{i-1})\right)^{1/2}\left(\sum_{i=1}^{n}(g'(\widetilde{t}_i))^2(t_i-t_{i-1})\right)^{1/2}.$$

As $\|\Pi\|\to 0$, by the square integrability of $f'$ and $g'$

$$\sum_{i=1}^{n}(f'(t_i^*))^2(t_i-t_{i-1})\to\int_0^T(f'(t))^2\,dt;$$
$$\sum_{i=1}^{n}(g'(\widetilde{t}_i))^2(t_i-t_{i-1})\to\int_0^T(g'(t))^2\,dt.$$

Since both integrals are finite, we conclude that $[f,g]_T=0$. This fact can be recorded informally as $df(t)dg(t)=f'(t)g'(t)dtdt=0$.

The above examples show that for "nice" functions (see Example 1.3 for conditions) the cross and quadratic variations are 0. But if $f$ and $g$ are "rough" (for example, $f(t)=g(t)=B(t)$ then the situation is different. Another point is that for random paths $f(t,\omega)$ and $g(t,\omega)$ the expressions under the limits in the definition of cross and quadratic variation are random variables. Therefore the limits can be understood in different senses (a.s., in $L^2$, in probability), and the existence of the limit might depend on the interpretation of the limit. We shall work with $L^2$ limits, since they happen to be the most natural ones for the stochastic calculus setting. Let us first recall the definition.

**Definition 1.4.** *Let $X_n$, $n\ge 1$, be a sequence of random variables on some probability space $(\Omega,\mathcal{F},P)$. We say that $X_n$ converges to random variable $X$ in $L^2$ (or in mean square) if*

$$\lim_{n\to\infty}E(X_n-X)^2=0.$$

Our first result will be about quadratic variation of Brownian motion.

**Theorem 1.5.** *Let $(B(t)_{t\ge 0}$ be a standard Brownian motion. Then $[B]_T=T$ for all $T\ge 0$ (we understand the limit in the $L^2$ sense).*

**Remark 1.6.** The statement $[B]_t=t$ for all $t\ge 0$ can be interpreted as follows: Brownian motion accumulates quadratic variation at rate 1 ($t'=1$). The fact that the quadratic variation for Brownian motion is non-random is an important property of Brownian motion. Just as in the case of "nice" functions, the above theorem gives a precise meaning to the informal expression $dB(t)dB(t)=dt$.

---

[2]Take discrete random variables $X$ and $Y$ such that $P(X=x_i)=P(Y=y_i)=1/n$, $i\in\{1,2,\dots,n\}$. Then the inequality $E(|XY|)\le(E(X^2))^{1/2}(E(Y^2))^{1/2}$ boils down to

$$\frac{1}{n}\sum_{i=1}^{n}|x_iy_i|\le\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}y_i^2\right)^{1/2}.$$

*Proof.* Let $0 = t_0 < t_1 < \cdots < t_n = T$ and $\delta_i := (B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})$. Then

$$\begin{aligned}\left(\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 - T\right)^2 &= \left(\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 - \sum_{i=1}^n (t_i - t_{i-1})\right)^2 \\ &= \left(\sum_{i=1}^n \delta_i\right)^2 = \sum_{i=1}^n \delta_i^2 + 2\sum_{i<j} \delta_i\delta_j. \end{aligned} \tag{1.2}$$

Since Brownian motion has independent increments, $E(\delta_i\delta_j) = E(\delta_i)E(\delta_j) = 0$ for $i \neq j$. Next,

$$\begin{aligned} E(\delta_i^2) &= E((B(t_i) - B(t_{i-1}))^4) - 2(t_i - t_{i-1})E((B(t_i) - B(t_{i-1})^2) + (t_i - t_{i-1})^2 \\ &= 3(t_i - t_{i-1})^2 - 2(t_i - t_{i-1})^2 + (t_i - t_{i-1})^2 = 2(t_i - t_{i-1})^2. \end{aligned}$$

Taking the expectations in (1.2) and substituting the above results we get

$$E\left(\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 - T\right)^2 = 2\sum_{i=1}^n (t_i - t_{i-1})^2 \le 2\|\Pi\|T \to 0 \quad \text{as} \quad \|\Pi\| \to 0. \qquad \square$$

**Exercise 1.** Let $(B(t))_{t\ge 0}$ be a standard Brownian motion. Show that the cross-variation of $B(t)$ and $t$ on $[0,T]$ is equal to zero. Interpret the limit in the $L^2$ sense. This exercise gives a meaning to the informal expression $dB(t)dt = 0$.

**Remark 1.7.** It is also easy to show that the cross-variation of $B(t)$ and $t$ on $[0,T]$ is equal to zero if we interpret the limit in a.s. sense, that is

$$P\left(\lim_{\|\Pi\|\to 0} \sum_{i=1}^n (B(t_i) - B(t_{i-1})(t_i - t_{i-1}) = 0\right) = 1.$$

Indeed,

$$\left|\sum_{i=1}^n (B(t_i) - B(t_{i-1})(t_i - t_{i-1})\right| \le \max_{i\in\{1,2,\dots,n\}} |B(t_i) - B(t_{i-1})| \sum_{i=1}^n (t_i - t_{i-1}) = T \max_{i\in\{1,2,\dots,n\}} |B(t_i) - B(t_{i-1})|.$$

Since Brownian motion is continuous on $[0,T]$, it is uniformly continuous on $[0,T]$, that is for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|B(t) - B(t')| < \varepsilon$ as soon as $|t - t'| < \delta$, $t, t' \in [0,T]$. Since $\|\Pi\| \to 0$, we just have to make $\|\Pi\| < \delta/T$ to ensure that $T \max_{i\in\{1,2,\dots,n\}} |B(t_i) - B(t_{i-1})| < \varepsilon$.

**Exercise 2** (Advanced)**.** Use Borel-Cantelli lemma, part 1 (see HW2), to show that if we consider only equally spaced partitions: $t_i^n = Ti/n$, $i \in \{0, 1, 2, \dots, n\}$, then

$$P\left(\lim_{n\to\infty} \sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n))^2 = T\right) = 1.$$

Hint: for $m \in \mathbb{N}$ consider sets $A_{m,n} = \left\{\left|\sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n)^2 - T\right| \ge m^{-1}\right\}$ and show that $\sum_{n=1}^\infty P(A_{m,n}) < \infty$. You will need the following version of Markov's inequality:

$$P(|X - E(X)| \ge a) \le E((X - E(X))^4)/a^4.$$

## 2. "Bare hands" approach: computing $\int_0^T B(t)\,dB(t)$

Let $f \in C([0,T])$. Recall that the Riemann integral is defined as a limit of Riemann sums as $\|\Pi\| \to 0$:

$$\int_0^T f(t)\,dt = \lim_{\|\Pi\|\to 0} \sum_{i=1}^n f(t_i^*)(t_i - t_{i-1}).$$

For $f \in C([0,T])$ this limits exists and does not depend on the choice of $t_i^* \in [t_{i-1}, t_i]$, $i \in \{0, 1, \dots, \}$. We can try to define the integral with respect to Brownian motion in a similar way: for a stochastic process $f(t)$, $t \geq 0$, with continuous sample paths (for example, take $f(t) = B(t)$, $t \geq 0$) set

$$\int_0^T f(t, \omega)\, dB(t) \overset{L^2}{=} \lim_{\|\Pi\| \to 0} \sum_{i=1}^n f(t_i^*)(B(t_i) - B(t_{i-1})), \tag{2.1}$$

where $t_i^* \in [t_{i-1}, t_i]$, $i \in \{0, 1, \dots, \}$, and the limit is understood in the $L^2$ sense as indicated with $\overset{L^2}{=}$. Then we shall find out that, unlike in the case of the Riemann integral, the limit might depend on the choice of $t_i^*$ (see Exercise 3 below). Natural choices seem to be (a) $t_i^* = t_{i-1}$ (the left endpoint of $[t_{i-1}, t_i]$); (b) $t_i^* = (t_{i-1} + t_i)/2$ (the midpoint); (c) $t_i^* = t_i$ (the right endpoint). Choice (a) corresponds to the Itô integral and choice (b) corresponds to the Stratonovich integral. Choice (c) is used rarely and, I believe, does not have a special name.

We shall work only with Itô stochastic integral. Stratonovich integral is also widely used[3], in particular, because the associated "chain rule" for Stratonovich integral (known as Itô's formula in the case of Itô integral) happens to be the same as in regular calculus (see again Exercise 3).

**Example 2.1.** Let us adopt the choice (a) and compute

$$\int_0^T B(t)\, dB(t) \overset{L^2}{=} \lim_{\|\Pi\| \to 0} \sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1})).$$

We shall use Theorem 1.5 which says that

$$\lim_{\|\Pi\| \to 0} E\left(\sum_{i=1}^n (B(t_i) - B(t_{i-1})^2 - T\right)^2 = 0.$$

Using the identity $(b-a)^2 = b^2 - a^2 - 2a(b-a)$ we can write

$$\begin{aligned}\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 &= \sum_{i=1}^n B^2(t_i) - \sum_{i=1}^n B^2(t_{i-1}) - 2\sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1})) \\ &= B^2(T) - B^2(0) - 2\sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1}).\end{aligned}$$

Therefore,

$$\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 - T = B^2(T) - T - 2\sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1}).$$

Since the left-hand side converges to 0 in the mean square sense, we conclude that the right-hand side converges to 0 in the same sense and

$$\lim_{\|\Pi\| \to 0} \sum_{i=1}^n B(t_{i-1})(B(t_i) - B(t_{i-1})) \overset{L^2}{=} \frac{1}{2}(B^2(T) - T).$$

We have shown that

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

The same result is often recorded in the differential form as

$$B(t)\, dB(t) = d\left(\frac{1}{2}B^2(t)\right) - \frac{1}{2}\, dt, \quad \text{and rearranging we get} \quad dB^2(t) = 2B(t)\, dB(t) + dt.$$

[3]but mostly not in finance, see p. 78 in Etheridge's book for reasons.

If $f(t)$ is a differentiable function then by the chain rule $df^2(t) = 2f(t)\,df(t)$. In the case of Brownian motion we get an additional term. Its appearance is due to the fact Brownian motion paths are "rough" and accumulate quadratic variation.

**Exercise 3.** Show that if we let $t_i^* = (t_i + t_{i-1})/2$, $i \in \{1, 2, \dots, n\}$, in (2.1) and understand the limit in the mean square sense then

$$\int_0^T B(t) \circ dB(t) = \frac{1}{2} B^2(T),$$

where "$\circ$" indicates that we are considering the Stratonovich version. The above equality can be written in the differential form as follows:

$$d\left(\frac{1}{2} B^2(t)\right) = B(t) \circ dB(t).$$

This explains why we say that for Stratonovich integral the chain rule looks the same as in regular calculus.

## 3. Itô integral for elementary integrands.

We would like to define the stochastic integral for a general stochastic process $\Delta(t)$, $t \in [0, T]$. The idea is to define the stochastic integral $\int_0^T \Delta(t)\,dB(t)$ as a limit (in the $L^2$ sense) of stochastic integrals $\int_0^T \Delta_n(t)\,dB(t)$ where $\Delta_n$, $n \in \mathbb{N}$, is a sequence of some elementary processes which approximates $\Delta$. As the first step, we shall introduce a class of elementary processes, define the Itô integral for them, and discuss its properties.

Let $\mathcal{F}(t)$, $t \geq 0$, be a filtration for Brownian motion $B(t)$, $t \geq 0$. Fix a partition $\Pi$ of $[0, T]$, $0 = t_0 < t_1 < \dots < t_n = T$, and suppose that

$$\Delta(t, \omega) = \sum_{j=1}^{n} \Delta_{j-1}(\omega) \mathbb{1}_{[t_{j-1}, t_j)}(t), \tag{3.1}$$

where $\Delta_{j-1}$ is $\mathcal{F}(t_{j-1})$-measurable square integrable random variable. Define $I(0) := 0$ and for $t \in [t_k, t_{k+1})$ let

$$I(t) := \sum_{j=1}^{k} \Delta_{j-1}(B(t_j) - B(t_{j-1})) + \Delta_k(B(t) - B(t_k)). \tag{3.2}$$

Observe that $I(t_j)$, $j \in \{0, 1, \dots, n\}$, is a discrete stochastic integral which we discussed in the refresher[4].

**Definition 3.1.** *Let $\Delta(t)$, $t \in [0, T]$, be a stochastic process of the form (3.1). Then for all $0 \leq r \leq t \leq T$ we set*

$$\int_r^t \Delta(s)\,dB(s) := I(t) - I(r),$$

*where $I(t)$, $t \in [0, T]$ is given by (3.2).*

**Theorem 3.2** (Properties of the stochastic integral (for elementary processes))**.** *Let $\Delta(t)$, $\Delta_1(t)$, $\Delta_2(t)$, $0 \leq t \leq T$ be elementary stochastic processes of the form (3.1) and $I(t)$, $0 \leq t \leq T$, be as in (3.2).*

- *(Path continuity) The stochastic process $I(t)$, $0 \leq t \leq T$, has continuous paths.*
- *(Linearity) Let $a, b \in \mathbb{R}$. Then $a\Delta_1(t) + b\Delta_2(t)$, $t \in [0, T]$, is an elementary stochastic process and*

$$\int_0^t a\Delta_1(s) + b\Delta_2(s)\,dB(s) = a\int_0^t \Delta_1(s)\,dB(s) + b\int_0^t \Delta_2(s)\,dB(s), \quad 0 \leq t \leq T.$$

[4]See Definition 5 from refresher lecture 5. There $M_j = B(t_j)$, $j \in \{0, 1, \dots, n\}$, $\mathcal{F}_j$ is $\mathcal{F}_{t_j}$, and $H_j = \Delta_{j-1}$, $i = \{1, 2 \dots, n\}$.

- *(Additivity) Let* $0 \le r \le t \le T$. *Then*
$$\int_0^t \Delta(s)\,dB(s) = \int_0^r \Delta(s)\,dB(s) + \int_r^t \Delta(s)\,dB(s).$$
- *(Martingale property) The process* $I(t)$, $0 \le t \le T$, *is an* $\mathcal{F}(t)$*-martingale. In particular,* $E(I(t)) = E(I(0)) = 0$.
- *(Itô's isometry) For all* $t \in [0,T]$
$$\mathrm{Var}(I(t)) = E\left[\left(\int_0^t \Delta(s)\,dB(s)\right)^2\right] = \int_0^t E(\Delta^2(s))\,ds.$$

*Proof.* Path continuity, linearity, and additivity immediately follow from the definition of the integral and continuity of Brownian motion paths. The proof of martingale property is very similar to the one given in refresher lecture 5 (Theorem 10)[5].

For the proof of Itô's isometry fix $t \in [0,T]$ and let $\delta_i = B(t_i) - B(t_{i-1}),\ i \in \{1,2,\dots,k,\},\ \delta_{k+1} = B(t) - B(t_k)$. Then
$$E\left[\left(\int_0^t \Delta(s)\,dB(s)\right)^2\right] = E\left[\left(\sum_{i=1}^{k+1} \Delta_{i-1}\delta_i\right)\left(\sum_{j=1}^{k+1} \Delta_{j-1}\delta_j\right)\right] = \sum_{i=1}^{k+1}\sum_{j=1}^{k+1} E(\delta_i\delta_j\Delta_{i-1}\Delta_{j-1}).$$
Consider three cases.

- If $i = j$ then by the fact that $\Delta_{i-1} \in \mathcal{F}_{i-1}$ and independence of $\delta_i$ from $\mathcal{F}(t_{i-1})$ we get
$$E(\delta_i^2\Delta_{i-1}^2) = E(\Delta_{i-1}^2)E(\delta_i^2).$$
- If $i < j$ then again by the same considerations as in the previous case (this time I decided to use conditional expectations)
$$E(\delta_i\delta_j\Delta_{i-1}\Delta_{j-1}) = E(E(\delta_i\delta_j\Delta_{i-1}\Delta_{j-1}|\,\mathcal{F}(t_{j-1}))) = E(\delta_i\Delta_{i-1}\Delta_{j-1}E(\delta_j\,|\,\mathcal{F}(t_{j-1}))) = 0.$$
- If $i > j$ then just exchange $i$ and $j$ everywhere in the previous case to get the same result.

We conclude that
$$\sum_{i=1}^{k+1} E(\delta_i^2\Delta_{i-1}^2) = \sum_{i=1}^{k+1} E(\Delta_{i-1}^2)(t_i - t_{i-1}) = \int_0^t E(\Delta^2(s))\,ds,$$
since $E(\Delta^2(t))$, $0 \le t \le T$, is a non-random elementary function and the last integral is just another way to write the Riemann sum for this function. □

**Exercise 4.** Give the details of the proof of the martingale property in Theorem 3.2.

**Theorem 3.3** (Quadratic variation of stochastic integral). *Let* $\Delta(t), \Delta_1(t), \Delta_2(t),\ 0 \le t \le T$, *be elementary stochastic processes and* $I(t),\ I_1(t),\ I_2(t)$ *be their Itô integrals defined as in* (3.2). *Then*
$$[I]_t := [I,I]_t = \int_0^t \Delta^2(s)\,ds, \quad 0 \le t \le T;$$
$$[I_1,I_2]_t = \int_0^t \Delta_1(s)\Delta_2(s)\,ds, \quad 0 \le t \le T.$$

**Remark 3.4.** Note that there are no expectations in the right-hand sides so that the quadratic variation of the stochastic integral and cross-variation of stochastic integrals are stochastic processes: for each $t \in [0,T]$ the integrals $\int_0^t \Delta^2(s)\,ds$ and $\int_0^t \Delta_1(s)\Delta_2(s)\,ds$ are random variables.

[5] In Theorem 10 we assumed that $H_j$ were bounded for each $j$ but this condition can be relaxed to square integrability with practically no changes.

*Sketch of the proof.* We shall only prove the formula for the quadratic variation. Cross-variation can be handled by polarization identity (1.1) and linearity of ordinary integrals.

Let $\Pi$ be an arbitrary partition of $[0, T]$. Then $\Delta(t)$ need not be constant on each interval of this new partition. Moreover, $t_j$ (see (3.1)) need not be included as points of the partition $\Pi$. But let us consider only partitions $\Pi$ which contain all $t_j$, $j \in \{1, 2, \dots, n\}$, as their points. Consider points of $\Pi$ which are in between $t_{j-1}$ and $t_j$:

$$t_{j-1} = s_0 < s_1 < \cdots < s_m = t_j, \quad m = m(j).$$

$\Delta(s)$ is a constant on $[t_{j-1}, t_j)$. Therefore,

$$\sum_{i=1}^{m} (I(s_i) - I(s_{i-1}))^2 = \Delta_{j-1}^2 \sum_{i=1}^{m} (B(s_i) - B(s_{i-1}))^2.$$

We know that the last sum converges in $L^2$ to $(t_j - t_{j-1})$ as $\|\Pi\| \to 0$. Moreover, the last sum is independent from $\Delta_{j-1}$ as $\Delta_{j-1}$ is $\mathcal{F}(t_{j-1})$-measurable. Thus,

$$\lim_{\|\Pi\| \to 0} \Delta_{j-1}^2 \sum_{i=1}^{m} (B(s_i) - B(s_{i-1}))^2 \overset{L^2}{=} \Delta_{j-1}^2 (t_j - t_{j-1}) = \int_{t_{j-1}}^{t_j} \Delta^2(s)\, ds. \tag{3.3}$$

This is true for each interval $[t_{j-1}, t_j)$, $j = 1, 2 \dots, k$, and for $[t_k, t]$. Let

$$a_j = \sum_{i=1}^{m(j)} (I(s_i) - I(s_{i-1}))^2 - \int_{t_{j-1}}^{t_j} \Delta^2(s)\, ds, \ j = 1, 2, \dots, k,$$

and agree to define $a_{k+1}$ similarly over the leftover interval $[t_k, t]$. Then

$$[I]_t - \int_0^t \Delta^2(s)\, ds = \sum_{j=1}^{k+1} a_j$$

By the Cauchy-Schwarz inequality,

$$\left( \sum_{j=1}^{k+1} a_j \right)^2 \le (k+1) \sum_{j=1}^{k+1} a_j^2.$$

Therefore,

$$E \left( \sum_{j=1}^{k+1} a_j \right)^2 \le (k+1) \sum_{j=1}^{k+1} E(a_j^2).$$

We have shown above that $E(a_j^2) \to 0$ as $\|\Pi\| \to 0$ for each $j = 1, 2, \dots, k+1$. Since $k + 1 \le n$ and $n$ is fixed (it is the number of "steps" in the function $\Delta$ on $[0, T]$), we obtain the result.

We omitted the proof of the fact that the restriction imposed on partitions $\Pi$ at the very beginning does not affect the existence or the value of the limits. □

Theorem 3.3 gives a precise meaning to the expression $dI(t)dI(t) = \Delta^2(t)\, dt$, for now only in the special case of elementary integrands. Notice that this can be formally obtained by "multiplication rules": $dI(t)dI(t) = \Delta(t)\, dB(t)\Delta(t)\, dB(t) = \Delta^2(t)\, dB(t)dB(t) = \Delta^2(t)\, dt$.

## 4. Itô integral for general integrands.

We shall only sketch the construction and state the main properties of the stochastic integral. These properties are extensions of those which we discussed in the previous section.

We have a probability space $(\Omega, \mathcal{F}, P)$, a Brownian motion on it with the associated filtration $\mathcal{F}(t) \subset \mathcal{F}$, $t \ge 0$. We shall always assume that the stochastic process $\Delta(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, satisfies the following two conditions:

(i) (Measurability) $\Delta : [0,T]\times\Omega \to \mathbb{R}$ is $\mathcal{B}_{[0,T]}\otimes\mathcal{F}$-measurable.[6]
(ii) (Adaptedness) $\Delta(t)$, $t\in[0,T]$, is adapted to the filtration $\mathcal{F}(t)$, $t\in[0,T]$.
(iii) (Square integrability) $\int_0^T E(\Delta^2(t))dt<\infty$.

Such processes can be approximated in the $L^2$ sense by elementary processes considered in the previous section, i.e. there is a sequence $\Delta^{(n)}=\Delta^{(n)}(t,\omega)$, $n\in\mathbb{N}$, of elementary processes such that

$$\lim_{n\to\infty}\int_0^T E(\Delta(t)-\Delta^{(n)}(t))^2dt=0. \tag{4.1}$$

The integrals $I_n(t):=\int_0^t \Delta^{(n)}(s)\,dB(s)$ are well-defined on $[0,T]$. We shall define $I(t)$, $t\in[0,T]$, as the limit of $I_n(t)$, $t\in[0,T]$, in the $L^2$ sense:

$$I(t)\overset{L^2}{=}\lim_{n\to\infty}I_n(t),\quad 0\le t\le T. \tag{4.2}$$

The existence of the limit needs justification as well as the fact that the limit does not depend on the approximating sequence (as long as (4.1) holds).[7]

**Remark 4.1.** To reconcile this definition with our “bare hands” approach to compute $\int_0^T B(t)\,dB(t)$ (see Section 2), we note that if we let $t_j=Tj/n$, $j\in\{0,1,\dots,n\}$, and define

$$\Delta^{(n)}(t,\omega)=\sum_{j=1}^n B(t_{j-1})\mathbb{1}_{[t_{j-1},t_j)}(t),\ j\in 1,2,\dots,n,$$

then we shall get an approximating sequence of elementary integrands satisfying (4.1) with $\Delta(t)=B(t)$. Thus, our computations in Section 2 amount exactly to showing (4.2) with $I(t)=\frac{1}{2}(B^2(T)-T)$.

We set

$$\int_r^t \Delta(s)\,dB(s):=I(t)-I(r),\quad 0\le r\le t\le T. \tag{4.3}$$

Properties of the stochastic integral are summarized (without a proof) in the following theorem.

**Theorem 4.2** (Properties of Itô integral)**.** *Let $\Delta,\Delta_1,\Delta_2:[0,T]\times\Omega\to\mathbb{R}$ satisfy the condition (i)-(iii) above. Then the stochastic integral defined in* (4.3) *has the following properties:*

- *(Adaptedness) $I(t)=\displaystyle\int_0^t\Delta(s)\,dB(s)$, $0\le t\le T$, is $\mathcal{F}(t)$-adapted.*
- *(Path continuity) The process $I(t)=\displaystyle\int_0^t\Delta(s)\,dB(s)$, $0\le t\le T$, has continuous paths.*
- *(Linearity) Let $a,b\in\mathbb{R}$. Then $a\Delta_1(t)+b\Delta_2(t)$, $t\in[0,T]$, satisfies (i)-(iii) and*
$$\int_0^t a\Delta_1(s)+b\Delta_2(s)\,dB(s)=a\int_0^t\Delta_1(s)\,dB(s)+b\int_0^t\Delta_2(s)\,dB(s),\quad 0\le t\le T.$$
- *(Additivity) Let $0\le r\le t\le T$. Then*
$$\int_0^t\Delta(s)\,dB(s)=\int_0^r\Delta(s)\,dB(s)+\int_r^t\Delta(s)\,dB(s).$$
- *(Martingale property) The process $I(t)$, $0\le t\le T$, is an $\mathcal{F}(t)$-martingale. In particular, $E(I(t))=E(I(0))=0$.*

---

[6] $\mathcal{B}_{[0,T]}\otimes\mathcal{F}$ is the product $\sigma$-algebra, i.e. the smallest $\sigma$-algebra which contains all sets of the form $B\times A$, where $B\in\mathcal{B}_{[0,T]}$, the Borel $\sigma$-algebra on $[0,T]$, and $A\in\mathcal{F}$.

[7] The details are beyond the scope of these notes and can be found in any standard text on stochastic calculus such as, for example, Øksendal, Stochastic Differential Equations, 2003, Chapter 3.

- *(Itô's isometry) For all* $t \in [0, T]$
$$\mathrm{Var}(I(t)) = E\left[\left(\int_0^t \Delta(s)\, dB(s)\right)^2\right] = \int_0^t E(\Delta^2(s))\, ds.$$
- *(Quadratic variation)* $[I]_t = \int_0^t \Delta^2(s)\, ds.$
- *(Cross-variation) Let* $I_i(t) = \int_0^t \Delta_i(s)\, dB(s)$, $i = 1, 2$, $t \in [0, T]$. *Then*
$$[I_1, I_2]_t = \int_0^t \Delta_1(s)\Delta_2(s)\, ds, \quad 0 \le t \le T.$$

**Example 4.3** (Application of Itô's isometry)**.** Let $I_j(t) = \int_0^t \Delta_j(s)\, dB(s)$, where $\Delta_j$ satisfies (i)-(iii), $j = 1, 2$. Find $\mathrm{Cov}(I_1(t), I_2(t))$.

Recall (or check) the following version of polarization identity: $xy = (1/4)((x+y)^2 - (x-y)^2)$. Applying this identity to $x = I_1(t)$ and $y = I_2(t)$ and taking the expectation we get

$$\mathrm{Cov}(I_1(t), I_2(t)) = E(I_1(t)I_2(t)) = \frac{1}{4} E((I_1(t) + I_2(t))^2 - (I_1(t) - I_2(t))^2). \tag{4.4}$$

By linearity and Itô's isometry,

$$E((I_1(t) \pm I_2(t))^2) = \int_0^t E(\Delta_1(s) \pm \Delta_2(s))^2\, ds.$$

Substituting in (4.4) and using the linearity of expectation and of the ordinary integral, we get

$$\mathrm{Cov}(I_1(t), I_2(t)) = \int_0^t E(\Delta_1(s)\Delta_2(s))\, ds = E([I_1, I_2]_t).$$

The last equality follows from the last part of Theorem 4.2.